G.S. ZHDANOV

CRYSTAL PHYSICS

G.S. ZHDANOV

CRYSTAL PHYSICS

PREFACE

Crystal physics is the study of the relationships between the composition, the atomic and electron structure and the various physical properties (electric, magnetic, mechanical, thermal etc.) of solids and in particular of crystals. The task of crystal physics includes the examination of such problems as how crystals are formed and grow (crystallization) and how they are broken down under the influence of various factors (melting, the effect of radiation etc.). It also includes the study of the behaviour of matter over a wide range of temperatures and pressures.

Crystal physics is at present one of the most important branches of science, and is of great practical utility. Modern technology is constantly posing problems concerning the creation of new materials with given properties. A short list of these might include materials possessing special mechanical (elasticity, solidity, rigidity, heat-resistance), electrical (semiconductors, superconductors, piezoelectrics, ferro- and antiferroelectrics), magnetic (ferro- and antiferromagnetics), optical (luminescence, crystals for optics in the infrared and ultraviolet) and other properties. In a number of cases materials are needed which possess a definite combination of properties, and this presents special difficulties. Together with engineers and technologists, physicists have an important part to play in the solution of these problems. The physicist has not only to work out physical methods for investigating and synthesising materials, but also to build up a logical microscopic theory of the crystalline state. A physicist who specialises in the solid state must have a wide-ranging mental outlook, so that he can work successfully not only in the established branches of physics but also in the new branches which are now just developing.

Crystal physics makes wide use of data on the atomic and electron structure of matter and the forces of interatomic and intermolecular interaction. These data are obtained with the aid of a number of important physical methods of investigation including structural (e.g. neutron diffraction), spectroscopic (e.g. radiospectroscopy, electron and nuclear resonance), radioactive tracers etc. Great interest has lately been centred on the use of strong nuclear radiations to produce structural damage in crystals and its application to the theory of real crystals, the class which comprises virtually all crystals with which one has to deal in practice. Even classical methods of physical and chemical analysis have not lost their importance, but have received further development and supplementation (electron microscope, ion projector etc.).

These methods together with quantum mechanics yield rich material to help with understanding the structure and properties of matter and the appreciation of the atomic and electron processes which take place in solids.

One of the purposes of writing this book was to make a more comprehensive collection of data on the atomic and electron structure of matter and the forces of interatomic and intermolecular interaction than is available in most textbooks and monographs on solid state physics. In this the author was helped by having been for a long time in charge of the crystal structure laboratory at the Karpov Physicochemical Institute. Undoubtedly the attitude to scientific enquiry of this laboratory has left its stamp on the book as a whole.

The nucleus of the book consisted of notes for lectures on solid state physics given by the author from 1948 onwards, at the Moscow Engineering Physics Institute to students specialising in the physics of metals, and from 1954 onwards in Moscow University to students in the Solid State Department of the Physics Faculty. Some chapters have problems appended, and in the appendix specimen seminar questions are given with the object of drawing the attention of students to individual problems in the course. As regards the great volume of material embraced by solid state physics, not all its divisions could be adequately treated. The book is provided with a bibliography of monographs, and also a partial bibliography of recent articles, with the aim of enabling anyone to broaden his knowledge if he becomes interested in questions which are treated in a very condensed manner in the book.

The manuscript has been revised by Academician N. V. Belov, Prof. V. I. Arkharov, Prof. K. P. Belov, and Prof. M. I. Zakharov; in addition, Prof. V. A. Fabrikant made a number of valuable observations, most of which the author has used. Considerable help in the preparation of the manuscript was given by the following research workers and postgraduate students: V. V. Zubyenko, V. Ya. Dudarev, S. P. Solovyev, N. V. Rannyev; also by laboratory workers S. Flyeyer, P. F. Kirsanova, L. D. Osipova, and by the Press editor, G. S. Goldenberg. To all the above-named persons, and also to his colleagues in the Crystal Structure Laboratory and Solid State Physics Department, with whom he had over a long period scientific discussions on a number of topics touched on in the book, the author wishes to express his thanks.

G. S. ZHDANOV

CONTENTS

PREFACE v

1 INTRODUCTION TO QUANTUM MECHANICS 1
1. Schrödinger's Equation for Particles with one Degree of Freedom 1
2. Schrödinger's Equation for a System of Particles 7
3. Particle in Fields of Constant Potential 10
4. Passage of a Particle through a Potential Barrier 17
5. Particle in a Rectangular Potential Well 20
6. The Harmonic Oscillator 27
7. Particle in a Spherically Symmetrical Field 30
8. Particle in a Periodic Field 39

2 ATOMIC STRUCTURE AND THE PERIODIC TABLE OF THE ELEMENTS 47
1. The Atomic Nucleus 47
2. Fundamental Properties of Atoms 53
3. The Hydrogen Atom 55
4. The Multi-electron Atom 64
5. The Periodic Law and the Periodic System of the Elements 68
6. The Electron Structure of Atoms of Different Groups of Elements 76

3 THE QUANTUM THEORY OF INTERATOMIC AND INTERMOLECULAR INTERACTION 82
1. Interaction between Atoms and Electrons 82
2. General Characteristics of the Forces of Interatomic Interaction 84
3. The Quantum Theory of Interatomic Interaction 87
4. Forces of Repulsion 98
5. Molecular Forces 101

4 THE BASIC TYPES OF CHEMICAL BOND 108
1. The Covalent Bond 108
2. The Metallic Bond 125
3. The Heteropolar Bond 129

5 THE SYMMETRY OF ATOMS, MOLECULES AND CRYSTALS 136
1. The Symmetry of Finite and Infinite Figures 136
2. Symmetry Groups and the Laws of their Formation 141
3. The Symmetry of Atoms and Molecules 145
4. The Symmetry of Crystals 147
5. The Symmetry of Atoms and the Types of Simplest Periodic Structures formed by them 155

6 THE STRUCTURE OF CRYSTALS: CRYSTAL CHEMISTRY 164
1. Crystal Chemistry of the Periodic System of the Elements 164
2. Structures of Solid Solutions 173
3. The Structures of Some Compounds 185

7 COHESIVE ENERGY 197
1. Cohesive Energy of Molecules and Crystals 197
2. Cohesive Energy of Molecular Crystals 200
3. Cohesive Energy of Ionic Crystals 203

8 THE ELECTRON THEORY OF CRYSTALS 212
1. Classical Theory of the Electron Gas 212
2. Bringing the Electron Model of a Metal into Agreement with Experimental Data 214
3. The Quantum Electron Theory of Crystals 217
4. The Band Theory of Crystals 223
5. The Theory of Electron Compounds 229

9 ELECTRICAL PROPERTIES OF CRYSTALS 233
1. Classification of Solids by their Electrical Properties 233
2. The Dielectric Constant and Polarisability 235
3. Crystalline Electrets and Pyroelectrics. Piezoelectrics 253
4. Ferro- and Antiferroelectrics 258
5. Electrical Conductivity of Crystals 296
6. Superconductivity 300

10 MAGNETIC PROPERTIES OF CRYSTALS 313
1. Basic Types of Magnet 313
2. Diamagnetism 317
3. Paramagnetism 321
4. Ferromagnetism 325
5. Antiferromagnetism and Magnetic Structure 329

11 PHASE EQUILIBRIUM AND PHASE TRANSFORMATIONS IN ALLOYS 333
1. Basic Conceptions 333
2. The Free Energy of a Disordered Solid Solution 335
3. Equilibrium States of an Alloy 338
4. Temperature Dependence of Solubility 340
5. Equilibrium Diagrams 342
6. Phase Transitions 347

12 ORDERING AND DISORDERING OF THE ATOMIC STRUCTURE OF ALLOYS 350
1. The Ordered Structure of Substitutional Solid Solutions 350
2. Theory of Long-Range Order 353
3. Short-Range Order 361

13 FORMATION AND GROWTH OF CRYSTALS 365
1. Basic Experimental Results Common to many Crystals 365
2. Thermodynamics of Crystal Formation 370
3. Kinetics of Crystallization 376

14 THERMAL PROPERTIES OF CRYSTALS 380
1. The Specific Heat of Crystals: Experimental Data 380
2. Principles of the Quantum Theory of the Specific Heat of Crystals 384
3. The Theory of Normal Vibrations of a Crystal 394
4. Thermal Conductivity of Crystals 402
5. Thermal Expansion of Crystals 411

15 SELF DIFFUSION AND DIFFUSION IN CRYSTALS 421
1. Basic Concepts and Experimental Results 421
2. The Mechanism of Diffusion in Crystals 424
3. Diffusion in Deformed Alloys 427
4. Diffusion in Solid Electrolytes 429

16 MECHANICAL PROPERTIES OF CRYSTALS 432
1. Elasticity, Brittleness and Plasticity of Crystals 432
2. Elasticity of Crystals 434
3. Cleavage and the Structure of Crystals 441
4. Plasticity of Crystals. Slip 445
5. Mechanical Twinning 458

APPENDIX 1. PROPERTIES OF THE ELEMENTS 461
TABLE I. First ionization potential of atoms, eV
II. Valencies
III. Density of the Solid Phase, gm/cm^3
IV. Heats of Sublimation, kcal/mole
V. Melting Points, °C
VI. Electrical Resistivities at 0° C, ohm. cm × 10^6
VII. Magnetic Susceptibilities ($\chi_A \times 10^6$) at Room Temperature

APPENDIX 2. SEMINAR PROBLEMS 468

APPENDIX 3. SOLUTIONS TO THE TEXT PROBLEMS 470

BIBLIOGRAPHY 491

INDEX OF AUTHORS 495

INDEX OF SUBJECTS 497

Chapter 1

INTRODUCTION TO QUANTUM MECHANICS

1. Schrödinger's Equation for Particles with one Degree of Freedom

Schrödinger's Equation. The Hamiltonian

Atoms, molecules and crystals are systems consisting of atomic nuclei and electrons. The motion of the electrons and atomic nuclei, and their interactions with one another, are subject to quantum laws. Therefore a concise account will first be given of some aspects of quantum mechanics which are essential for understanding those problems to be dealt with in later sections.

Mathematically, the behaviour of quantum systems is described by Schrödinger's equation

$$H\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \qquad \ldots [1.1]$$

Here h is Planck's constant;

$$\Psi = \Psi(x_1, y_1, z_1;\ x_2, y_2, z_2; \ldots, t)$$

is the wave function, depending on the coordinates of the particles x_1, y_1, z_1; $x_2, y_2, z_2, \ldots x_k, y_k, z_k$, and on the time t; $\boldsymbol{H}$ is the Hamiltonian operator, which is obtained from the Hamiltonian of classical mechanics by the following rules: The Hamiltonian is equal to the sum of the kinetic (T) and potential (U) energies:

$$H = T + U = E, \qquad \ldots [1.2]$$

that is, to the total energy (E) of the system. For conservative systems formula [*2*] which can be written in the form

$$H = E, \qquad \ldots [1.3]$$

expresses the law of conservation of energy. The kinetic energy of the particles in the system may be expressed in terms of the momenta p_k. In this case the Hamiltonian appears as a function of the coordinates and momenta of the particles. For a system with one degree of freedom (a particle of mass m, moving along the x-axis),

$$H(p_x, x) = \frac{1}{2m}p_x^2 + U(x) = E. \qquad \ldots [1.4]$$

To obtain the Hamiltonian operator, the momenta p_k in the classical Hamiltonian must be replaced by the differential operators

$$p_k = \frac{h}{2\pi i}\frac{\partial}{\partial x_k}. \qquad \ldots [1.5]$$

Consequently,

$$H = H\left(\frac{h}{2\pi i}\frac{\partial}{\partial x_k}, x_k; \ldots\right). \qquad \ldots [1.6]$$

The squares of the momenta p_k^2 are correspondingly replaced by operators involving second partial derivatives:

$$p_k^2 = \left(\frac{h}{2\pi i}\right)^2 \frac{\partial^2}{\partial x_k^2} = -\frac{h^2}{4\pi^2}\frac{\partial^2}{\partial x_k^2}. \qquad \ldots [1.7]$$

The total energy in [*2*] must be replaced by a differential operator of the form

$$E \rightarrow -\frac{h}{2\pi i}\frac{\partial}{\partial t}. \qquad \ldots [1.8]$$

Schrödinger's equation is obtained from the equation [*3*] of classical mechanics by replacing the Hamiltonian and the total energy by the differential operators [*6*] and [*8*] and applying these operators to the wave function Ψ of the system. In operator form Schrödinger's equation takes the form

$$H\Psi = E\Psi. \qquad \ldots [1.9]$$

Equation [*9*], equivalent to equation [*1*], is called *Schrödinger's time-dependent equation.* For a particle with one degree of freedom this equation, when written out in full, takes the form

$$-\frac{h^2}{8\pi^2 m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + U(x)\Psi(x,t) = \frac{h}{2\pi i}\frac{\partial\Psi(x,t)}{\partial t}. \qquad \ldots [1.10]$$

In mathematical physics we meet three fundamental types of linear differential equations involving partial derivatives of the second order: Poisson's equation

$$\nabla^2 v = 4\pi\rho, \qquad \ldots [1.11]$$

of which a particular case is Laplace's equation

$$\nabla^2 v = 0; \qquad \ldots [1.12]$$

the equation of the form

$$\nabla^2 v = \frac{1}{k}\frac{\partial v}{\partial t}; \qquad \ldots [1.13]$$

and, finally, the equation

$$\nabla^2 v = \frac{1}{c^2}\frac{\partial^2 v}{\partial t^2}. \qquad \ldots [1.14]$$

In these equations ∇^2 is Laplace's differential operator, which in Cartesian coordinates takes the form

$$\nabla^2 = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2};$$

v is some function of the coordinates and the time. Equations [*11*] and [*12*] describe equilibrium states (potential of a gravitational field or potentials of electric and magnetic fields with or without sources of density ρ). Equation [*13*] describes processes leading to equilibrium: heat conduction (trend to thermal equilibrium), diffusion (which evens out the density of matter), internal friction (which evens out momenta) and so forth. Equation [*14*] describes oscillatory and wave processes, and is the fundamental differential equation of acoustics, optics and the electrodynamics of variable fields. The identity of the left hand sides of equations [*11*]-[*14*], which describe different physical phenomena, stems from the symmetry properties of a homogeneous isotropic medium. These require that the equations be invariant with respect to translation and rotation of the coordinate axes, and are seen, in particular, in the symmetrical form of Laplace's operator. The differential form of the equations follows from the concept of near neighbour interaction, the main point of which is that only neighbouring elements of the medium can exert any influence on one another.

Schrödinger's equation is also a linear differential equation containing partial derivatives of the second order; using Laplace's operator we can write it in the form

$$\nabla^2\Psi-\frac{8\pi^2 m}{h^2}U\Psi = \frac{8\pi m}{hi}\frac{\partial\Psi}{\partial t}. \qquad \ldots [1.15]$$

Comparison with equations [*11*]-[*14*] shows that Schrödinger's equation can be regarded as a generalised equation describing the trend to equilibrium. The difference between it and equation [*13*] lies in the fact that Schrödinger's equation has an imaginary factor on its right hand side. This leads to the essential difference in the nature of solutions of equations [*13*] and [*15*]: the usual diffusion equation does not have solutions which are periodic in time, whereas Schrödinger's equation does have periodic solutions. Moreover the left hand side of [*15*] contains a second term which vanishes in the absence of external fields, when $U = 0$.

Separation of the Variables

Suppose that the solution of equation [*15*] can be represented as the produce of two functions, one of which depends only on the coordinates, and the other only on the time:

$$\Psi(x, t) = \psi(x)\,\phi(t). \qquad \ldots [1.16]$$

Substituting [*16*] in [*10*], we obtain

$$\frac{1}{\psi(x)}\left[-\frac{h^2}{8\pi^2 m}\frac{d^2\psi(x)}{dx^2}+U(x)\psi(x)\right] = -\frac{h}{2\pi i}\frac{1}{\phi(t)}\frac{\partial\phi(t)}{dt}. \qquad \ldots [1.17]$$

The left hand side of [*17*] is a function of x only, and the right hand of t only. The variables x and t being independent, the only possible conclusion is that both sides of equation [*17*] are equal to a constant, which we shall call E. This quantity has the dimensions of energy. Equation [*17*] breaks down into two equations, one of which depends only on the time, and the other only on the coordinates.

$$\frac{d\phi(t)}{dt} = -\frac{2\pi i}{h} E\phi(t), \qquad \ldots [1.18]$$

$$\frac{d^2\psi(x)}{dx^2} + \frac{8\pi^2 m}{h^2} [E - U(x)]\psi(x) = 0. \qquad \ldots [1.19]$$

Equation [*18*] is easily integrated, and its solution is of the form

$$\phi_n(t) = A_n e^{-2\pi i \frac{E_n}{h} t} = A_n e^{-2\pi i \nu_n t} = A_n e^{-i\omega_n t}. \qquad \ldots [1.20]$$

Here E_n denotes one of the possible values of the parameter E, for, as will be shown later, equation [*19*] has acceptable solutions only for strictly determined values of this parameter: $E_1, E_2, \ldots E_n$, called the *eigenvalues*. The ratio

$$\frac{E_n}{h} = \nu_n = \frac{\omega_n}{2\pi} \quad (\omega_n = 2\pi\nu_n) \qquad \ldots [1.21]$$

has the dimensions of frequency. Formula [*20*] shows that when $E = \text{const}$, the solutions of Schrödinger's equation are sinusoidal oscillations periodic in time. Schrödinger's equation, unlike the classical wave equation, is unsymmetrical with respect to the variables x and t. So functions like the travelling waves $F(x, t) = F(x \pm ct)$, which are solutions of the wave equation, will not in general be solutions of Schrödinger's equation.

In what follows our main interest will centre on equation [*19*], which is known as *Schrödinger's time-independent equation.* This equation determines the space distribution of the wave function amplitude. Its solutions are the set of eigenfunctions

$$\psi_1(x), \psi_2(x), \ldots, \psi_n(x). \qquad \ldots [1.22]$$

[*19*] is a linear equation. The sum of any set of particular solutions of it is also a solution of this equation. The general solution is the sum of all the particular solutions

$$\Psi(x,t) = \sum_n a_n \Psi_n(x,t) = \sum_n a_n \Psi_n(x) e^{-2\pi i \frac{E_n}{h} t}, \qquad \ldots [1.23]$$

where the a_n are constant coefficients whose values are determined from the initial and boundary conditions of the problem. If the spectrum of frequencies $\nu_1, \nu_2, \ldots$ is continuous, then the summation sign in equation [*23*] is replaced by an integral sign.

Initial and Boundary Conditions

Initial conditions can be given in quantum mechanics in exactly the same way as is usual in the solution of classical problems expressible in the form of differential equations, that is, the value of the required function is given at a certain moment of time, for example

$$t = 0: \quad \Psi_{t=0} = \Psi(x, 0). \qquad \ldots [1.24]$$

This function gives the amplitude distribution in space at the initial instant of time.

In the investigation of stationary states of quantum systems (which will be our first interest) the nature of the solution does not depend on the initial conditions.

As regards *boundary conditions*, problems in quantum mechanics differ fundamentally from the majority of classical problems, since quantum systems, for example atoms, do not have boundaries in the ordinary sense. In such systems the wave functions have non-zero values in a bounded region of space beyond which they gradually fade away. The wave functions ψ_n are subject to the general limitations stemming from the physical conditions which are satisfied not only by wave functions, but also by functions of mathematical physics in general. These conditions are three in number: the function must be continuous, single-valued and bounded in every region of space. These conditions turn out to be sufficient for the solution of quantum-mechanical problems. This means that from among the multiplicity of mathematically possible solutions they enable us to choose those functions which describe the behaviour of the given quantum system.

The Physical Meaning of the Wave Function

The product of the wave function and its complex conjugate, $w = \Psi\Psi^*$, gives a real function defined for every region of variation of the variables. This function is proportional to the probability that at time t the system has the configuration described by the corresponding values of the arguments of the wave function. For a system with one degree of freedom the product

$$w(x,t) = \Psi(x,t)\Psi^*(x,t) \qquad \ldots [1.25]$$

is equal to the probability that at time t the particle lies in the region between x and $x+dx$. For the problems considered earlier, the number of particles in the system does not vary, and as a result the normalisation condition for the wave functions is

$$\int_{-\infty}^{+\infty} \Psi(x,t)\Psi^*(x,t)\,dx = 1, \qquad \ldots [1.26]$$

since the total probability that the particle will be somewhere in the interval $-\infty < x < +\infty$ is unity. The normalisation condition is applicable also to the amplitude part of the wave function:

$$\int_{-\infty}^{+\infty} \psi(x)\psi^*(x)\,dx = 1. \qquad \ldots [1.27]$$

Wave functions describing different states ψ_m^* and ψ_n^* satisfy the orthogonality condition

$$\int_{-\infty}^{+\infty} \psi_m^*(x)\psi_n(x)\,dx = 0, \qquad \ldots [1.28]$$

which stems from the properties of the solution of Schrödinger's equation. The proof of the orthogonality of the wave functions for the general case of a system of particles is given in § 2.

Degenerate States of Quantum Systems. Hybrid Wave Functions and their Orthogonalisation

If $E_n = E_m$ and the corresponding wave functions ψ_n and ψ_m are different and linearly independent, then to one and the same value of the energy E_n there correspond, not one, but several different states of the system,

$$\psi_{n1}, \psi_{n2}, \ldots \psi_{nk}, \ldots \psi_{nK}, \qquad \ldots [1.29]$$

which are called degenerate states. The number of different states, $k_{\max} = K$, belonging to a single energy level is called the *multiplicity* of the degeneracy. Degenerate states of quantum systems are frequently encountered. For example, all excited states of the hydrogen atom for a given $n > 1$ are degenerate. Another example of degeneracy is a system consisting of identical free atoms. Since any linear combination of solutions of a linear differential equation is also a solution of that equation, degenerate states can be described by different equivalent systems of wave functions. For instance, from two degenerate functions ψ_{n_1} and ψ_{n_2} we can form two new wave functions

$$\psi'_{n_1} = a_1\psi_{n_1} + a_2\psi_{n_2} \text{ and } \psi'_{n_2} = b_1\psi_{n_1} + b_2\psi_{n_2}, \qquad \ldots [1.30]$$

where a_1, a_2, b_1, b_2 are arbitrary constants which determine the degree of participation of each of the initial wave functions in the formation of *mixed* or, as they are often called, *hybrid* wave functions. If all the coefficients in [*30*] except one, say a_1, are zero, then out of all the possible degenerate states, the quantum system is in one particular state, the one described by the function ψ_{n_1}. If the initial system [*29*] of degenerate wave functions is not orthogonal, then by a suitable choice of hybrid functions it is always possible to reduce it to an orthogonal system. Thus for the functions [*30*] the orthogonality condition will take the form

$$\int \psi'_{n_1}\psi'_{n_2} d\tau = a_1 b_1 \int |\psi_{n_1}|^2 d\tau + (a_2 b_1 + a_1 b_2) \int \psi_{n_1}\psi_{n_2} d\tau$$
$$+ a_2 b_2 \int |\psi_{n_2}|^2 d\tau = 0;$$

there are many methods of satisfying this condition by a suitable choice of the coefficients a_1, a_2, b_1, b_2.

For the system of standardised and normalised functions expressed in formulae [*23*] and [*25*], the coefficients satisfy the relationship

$$\sum_n a_n a_n^* = 1. \qquad \ldots [1.31]$$

Stationary States

The probability distribution function $\Psi^*\Psi$ of a system in the state described by the wave functions [*23*] and [*25*] has the form

$$\Psi^*(x, t)\Psi(x, t) = \sum_n a_n^* a_n \psi_n^*(x)\psi_n(x) +$$
$$+ \sum_m \sum_n{}' a_m^* a_n \psi_m^*(x)\psi_n(x) e^{2\pi i \frac{(E_m - E_n)}{h} t}. \qquad \ldots [1.32]$$

In the double summation the prime denotes that the summation is carried out only for values of the indices such that $n \neq m$. In general, the probability function and consequently the properties of the system depend on the time, since the time enters into the second summation, which contains terms corresponding to transitions between states with different energy levels. If the system is in one of the states with energy E_n, then all the coefficients in [*23*] vanish except one: $a_n \neq 0$. In this case the double sum in [*30*] becomes zero and the probability function $\Psi^*\Psi$ becomes independent of the time and describes the stationary states of the system.

The mean value of any quantity which is a function $F(x)$ of the coordinates of the system is given by the formula

$$\bar{F} = \int_{-\infty}^{+\infty} \Psi^*(x,t)\Psi(x,t)F(x)\,dx. \qquad \ldots [1.33]$$

The mean value of dynamical quantities described by functions of the type $g(p_x, x)$, depending only on the momenta p_x, is given by the formula

$$\bar{g} = \int_{-\infty}^{+\infty} \Psi^*(x,t)\,g\left(\frac{h}{2\pi i}\frac{\partial}{\partial x}, x\right)\Psi(x,t)\,dx, \qquad \ldots [1.34]$$

in which the dynamical function $g(p_x, x)$ is replaced by the corresponding operator $\boldsymbol{g}$, which operates on the wave function $\Psi(x, t)$.

2. Schrödinger's Equation for a System of Particles

The Time-dependent Equation

Let us consider a system of N particles with masses $m_1, m_2, \ldots m_j, \ldots m_N$, moving in three-dimensional space. The potential energy U of the system depends in the general case on the $3N$ Cartesian coordinates of the particles, $x_1, y_1, z_1, \ldots x_j, y_j, z_j, \ldots x_N, y_N, z_N$ and the time t: $U = U(x_1, \ldots, z_N, t)$.

We shall consider conservative systems, for which the potential energy depends only on the configuration of the system and not on the time: $U = U(x_1, \ldots, z_N)$. The behaviour of such a system is described by Schrödinger's equation [*1*], in which the wave function $\Psi = \Psi(x_1, \ldots, z_N, t)$ is a function of the $3N$ coordinates of the system and the time. The operator $\boldsymbol{H}$ is the total Hamiltonian of the system [*6*], which is obtained from the Hamiltonian function by the rule given in formula [*7*]. In expanded form for a system of particles, equation [*1*] becomes

$$-\frac{h^2}{8\pi^2}\sum_{j=1}^{N}\frac{1}{m_j}\left(\frac{\partial^2\Psi}{\partial x_j^2}+\frac{\partial^2\Psi}{\partial y_j^2}+\frac{\partial^2\Psi}{\partial z_j^2}\right)+U(x_1,\ldots z_N)\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}.$$

Using the symbol for Laplace's operator, we can write this equation in the form

$$-\frac{h^2}{8\pi^2}\sum_{j=1}^{N}\frac{1}{m_j}\nabla_j^2\Psi+U\Psi = -\frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}. \qquad \ldots [1.35]$$

The Equation for the Amplitudes

As in the case of one particle, by putting

$$\Psi(x_1,\ldots z_N, t) = \Psi(x_1,\ldots, z_N)\,\phi(t) \qquad \ldots [1.36]$$

we can separate the variables and obtain two equations:

$$\frac{\partial\phi(t)}{dt} = -\frac{2\pi i}{h}E\phi(t), \qquad \ldots [1.37]$$

$$\sum_{j=1}^{N}\frac{1}{m_j}\nabla_j^2\psi(x_1,\ldots,z_N)+\frac{8\pi^2}{h^2}[E-U(x_1,\ldots,z_N)]\psi(x_1,\ldots,z_N) = 0, \quad \ldots [1.38]$$

where E is a constant parameter. For a single particle, equation [*37*] is the same as equation [*18*] for one particle, and its solution is [*20*]. The amplitude part of the wave function, $\psi(x_1,\ldots,x_N)$, satisfies the general conditions of continuity, single-valuedness and finiteness in every region of variation, from $-\infty$ to $+\infty$, of each of the $3N$ Cartesian coordinates. Under these conditions, as in the one-dimensional case, equation [*38*] is soluble only for strictly defined values of the parameter E (the energy of the system), which may form either a discrete or a continuous spectrum or a combination of both forms of spectra.

Corresponding to the total number of degrees of freedom, the state of the system is characterised by $3N$ integral quantum numbers $n_1,\ldots,n_{3N}$. These numbers uniquely determine both the wave function $\psi_{n_1,\ldots,n_{3N}}$ and the energy of the system $E_{n_1,\ldots,n_{3N}}$. For brevity the set of quantum numbers will be denoted by a single symbol n and ψ_n and E_n used as abbreviations. As in the one-dimensional case, the system may be degenerate, that is, to a single energy value E_n there may correspond a set of different states described by functions $\psi_n, \psi_n', \ldots$.

A particular solution of equation [*35*] has the form

$$\Psi_n(x_1,\ldots,z_N,t) = \psi_n(x_1,\ldots,z_N)e^{-2\pi i\frac{E_n}{h}t}. \qquad \ldots [1.39]$$

The general solution is the sum of all the particular solutions

$$\Psi(x_1,\ldots,z_N,t) = \sum_n a_n\psi_n(x_1,\ldots,z_N)e^{-2\pi i\frac{E_n}{h}t}, \qquad \ldots [1.40]$$

where the constants a_n are determined by the conditions of the problem.

The Physical Meaning of the Wave Function. Normalisation and Orthogonality

Besides the function Ψ, the state of the system is also described by its complex conjugate function Ψ^*, which is the solution of the equation conjugate to [*35*]. The real function $\Psi\Psi^*$ is a measure of the probability of the system being in a particular configuration. The quantity

$$\Psi_n^*(x_1,\ldots,z_N,t)\,\Psi_n(x_1,\ldots,z_N,t)\,dx_1\,dy_1\,dz_1\ldots dx_j\,dy_j\,dz_j\ldots,dx_N\,dy_N\,dz_N$$

expresses the probability that at time t the particle j is in the volume element bounded by the coordinate values x_j, x_j+dx_j; y_j, y_j+dy_j; z_j, z_j+dz_j, where j takes the values $1, 2, \ldots, N$, corresponding to the various particles. The wave functions must be *normalised* to unity for each particle:

$$\int\Psi_n^*\Psi_n\,d\tau = 1, \qquad \ldots [1.41]$$

where $d\tau$ is the volume element in the configuration space. For a normalised system of functions the coefficients a_n in [*40*] satisfy the relationship

$$\sum_n a_n^* a_n = 1.$$

The normalisation condition holds also for the amplitude parts of the wave functions:

$$\int \psi_n^*(x_1, \ldots, z_N)\psi_n(x_1, \ldots, z_N)\,d\tau = 1. \qquad \ldots [1.42]$$

The wave functions also satisfy the orthogonality conditions

$$\int \psi_m^* \psi_n \, d\tau = 0 \quad (m \neq n). \qquad \ldots [1.43]$$

To prove [*43*] we substitute into equation [*38*] its solutions ψ_m^* and ψ_n obtaining

$$\sum_{j=1}^{N} \frac{1}{m_j} \nabla_j^2 \psi_m^* + \frac{8\pi^2}{h^2}(E_m - U)\psi_m = 0$$

and

$$\sum_{j=1}^{N} \frac{1}{m_j} \nabla_j^2 \psi_n + \frac{8\pi^2}{h^2}(E_n - U)\psi_n = 0.$$

Multiplying the first of these equations by ψ_n and the second by ψ_m^*, and then subtracting the second from the first, we obtain

$$\sum_{i=1}^{N} \frac{1}{m_i}(\psi_m \nabla_i^2 \psi_n - \psi_n \nabla_i^2 \psi_m) + \frac{8\pi^2}{h^2}(E_n - E_m)\psi_n \psi_m = 0.$$

Integrate both terms of this equation throughout the whole of the configuration space:

$$\frac{8\pi^2}{h^2}(E_n - E_m)\int \psi_n \psi_m \, d\tau = -\sum_{i=1}^{N} \frac{1}{m_i} \int (\psi_m \nabla_i^2 \psi_n - \psi_n \nabla_i^2 \psi_m)\,d\tau. \qquad \ldots [1.44]$$

The R.H.S. of the equation obtained is then written in the coordinate form

$$\sum_{j}^{3N} \frac{1}{m_i} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \left(\psi_m \frac{\partial^2 \psi_n}{\partial q_j^2} - \psi_n \frac{\partial^2 \psi_m}{\partial q_j^2}\right) dq_1 dq_2 \ldots dq_{3N}.$$

The expression under the integral sign may be represented in the form

$$\psi_m \frac{\partial^2 \psi_n}{\partial q_j^2} - \psi_n \frac{\partial^2 \psi_m}{\partial q_j^2} = \frac{\partial}{\partial q_j}\left(\psi_m \frac{\partial \psi_n}{\partial q_j} - \psi_n \frac{\partial \psi_m}{\partial q_j}\right).$$

Integrating both sides of this identity with respect to the coordinate concerned, we find:

$$\int_{-\infty}^{+\infty} \left(\psi_m \frac{\partial^2 \psi_n}{\partial q_j^2} - \psi_n \frac{\partial^2 \psi_m}{\partial q_j^2}\right) dq_j = \left[\psi_m \frac{\partial \psi_n}{\partial q_j} - \psi_n \frac{\partial \psi_m}{\partial q_j}\right]_{-\infty}^{+\infty} = 0,$$

since $\psi_m, \psi_n \to 0$ at infinity. Consequently, we have from [*44*]

$$\frac{8\pi^2}{h^2}(E_n - E_m)\int \psi_n\psi_m d\tau = 0.$$

Provided $(E_n - E_m) \neq 0$, this gives us the orthogonality condition [*43*].

3. Particle in Fields of Constant Potential

The Free Particle

Let a particle of mass m move along the x-axis in a field of constant potential $U = \text{const}$. Schrödinger's equation in this case will be written in the form

$$\psi''(x) + k^2\psi(x) = 0, \qquad \dots [1.45]$$

where k is a constant parameter:

$$k = \frac{2\pi}{h}\sqrt{2m(E-U)}. \qquad \dots [1.46]$$

Particular solutions of equation [*45*] are functions of the type

$$\psi(x) = Ae^{\pm ikx}, \qquad \dots [1.47]$$

where A is a constant coefficient. The general solution is the sum of the two particular solutions:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}. \qquad \dots [1.48]$$

For a free particle the kinetic energy is

$$T = E - U = \tfrac{1}{2}mv^2,$$

making the sign of the expression under the square-root sign in [*46*] positive, and so the parameter k real. Equation [*45*] in this case is the same as the equation for harmonic oscillations, and its solution [*48*] is the superposition of two sinusoidal waves. Multiplying both sides of [*48*] by the time factor [*20*], we obtain the wave function for the particle as a function of the coordinate and the time:

$$\Psi(x,t) = Ae^{i(kx-\omega t)} + Be^{-i(kx+\omega t)}. \qquad \dots [1.49]$$

The general solution of [*49*] is a superposition of two running waves with the same frequency ω and the same wavelength λ, moving in opposite directions. For a particle moving in the direction of the positive x-axis, we must put $B = 0$, and for a particle moving in the opposite direction $A = 0$. The quantity k, which in the general case is vectorial, is called the wave vector, and its direction is parallel to the velocity v of the particle and coincides with the direction of the normal to the wave front. The length of the wave vector is inversely proportional to the de Broglie wavelength:

$$k = \frac{2\pi}{\lambda} \qquad \dots [1.50]$$

and, as can easily be seen from [46], directly proportional to the momentum $p = mv$ of the particle:

$$k = \frac{2\pi p}{h}. \qquad \ldots [1.51]$$

[50] and [51] imply de Broglie's law $\lambda = h/p$ (this should not be regarded as a proof of de Broglie's law, as the latter is the basis of Schrödinger's equation). From [46] we obtain the connection between the particle energy and the length of the wave vector (for simplicity we assume $U = 0$):

$$E = \frac{h^2}{8\pi^2 m} k^2. \qquad \ldots [1.52]$$

This function represents a parabola (Fig. 1). If the condition $k^2 > 0$ is observed, equation [46] has solutions for any value of the parameter k. This means that the energy spectrum of a free particle is continuous and the energy E can take any value. The probability of finding the particle in the section of its trajectory from x to $x+dx$ is given by the square of the modulus of the wave function [47]:

$$dw = \psi\psi^* dx = A^2 dx.$$

This probability is proportional to the length of the relevant section of the trajectory and has a constant value along the whole trajectory.

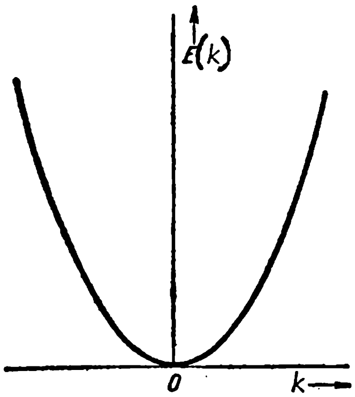


FIG. 1. Dependence of the energy E of a free particle on the length k of the wave vector.

Phase and Group Velocities of de Broglie Waves

From the equation [49] for the de Broglie wave we shall determine its velocity of propagation, which we shall call u. We shall find the velocity of translation of a section of the wave with uniform phase from the condition $kx - \omega t = \text{const}$. The latter gives

$$u = \frac{dx}{dt} = \frac{\omega}{k}. \qquad \ldots [1.53]$$

The frequency ω and the modulus k of the wave vector are connected, as formulae [21] and [52] show, with the particle energy. This enables us to express the phase velocity purely in terms of the wave vector, or by formula [50], of the de Broglie wavelength:

$$u = \frac{h}{2m} \cdot \frac{1}{\lambda}. \qquad \ldots [1.54]$$

Formula [54] shows that the velocity of propagation of de Broglie waves depends on the wavelength, and consequently these waves experience dispersion. The concept of a strictly monochromatic wave is an abstraction in physics. The monochromatic waves which we encounter under real conditions represent *groups* of waves with slightly differing wave vectors $k+\Delta k$, that is, with almost identical wavelengths and directions of propagation and occupying a bounded region in space. Because of dispersion and interference the group propagation velocity V (for example, the translation velocity of the intensity maximum) is different from the phase velocities of the component waves.

We shall find the *group velocity* V for the simplest group, which is formed by superposing two waves with neighbouring frequencies ω and ω', wave vectors k and k' and identical amplitudes. The resulting wave is described by the equation

$$\psi(x, t) = \cos(kx-\omega t) + \cos(k'x-\omega' t) =$$

$$= 2\cos\left(\frac{k-k'}{2}x-\frac{\omega-\omega'}{2}t\right)\cos\left(\frac{k+k'}{2}x-\frac{\omega+\omega'}{2}t\right)$$

and represents beats. The first trigonometrical factor varies slowly as compared with the second and can be regarded as the amplitude of the wave. The second factor describes the phase of the wave. The phase velocity is

$$u = \lim_{\omega'\to\omega}\frac{\omega+\omega'}{k+k'} = \frac{\omega}{k},$$

which is the same as [*53*]. The group velocity, which gives the translation velocity of the maximum of the beat, is

$$V = \lim_{\omega'\to\omega}\frac{\omega-\omega'}{k-k'} = \frac{d\omega}{dk}. \qquad \ldots [1.55]$$

In the case of dispersion $\frac{d\omega}{dk} \neq \frac{\omega}{k}$ and $V \neq u$. Formula [*55*] is also true in the general case, when the group is a superposition of a large number of sinusoidal waves and is described by a Fourier integral.

For de Broglie waves

$$\omega = 2\pi\frac{mv^2}{2h}, \quad k = 2\pi\frac{mv}{h}. \qquad \ldots [1.56]$$

Hence

$$V = \frac{d\omega}{dk} = \frac{\frac{d\omega}{dv}}{\frac{dk}{dv}} = v. \qquad \ldots [1.57]$$

Thus the group velocity of de Broglie waves is equal to the mechanical velocity of the corresponding particle.

Passage of a Particle across a Potential Step

Let a particle be moving through a region in which the potential U varies. We shall consider the simplest case of the motion of a particle in two constant fields meeting one another at the point $x = a$ and having different potentials U_1 and U_2. These fields form a potential step of height $U = U_2 - U_1$.

According as the particle energy is greater or less than the height of the step, we must distinguish between low ($E > U$) and high ($E < U$) potential steps (Fig. 2). In each of the fields the particle moves freely and its motion is described by Schrödinger's equation [*45*] with the parameter k constant for each field. At the boundary where the step occurs the value of k changes discontinuously, and so the wave function $\psi(x)$ changes. We first write down Schrödinger's equation for both fields:

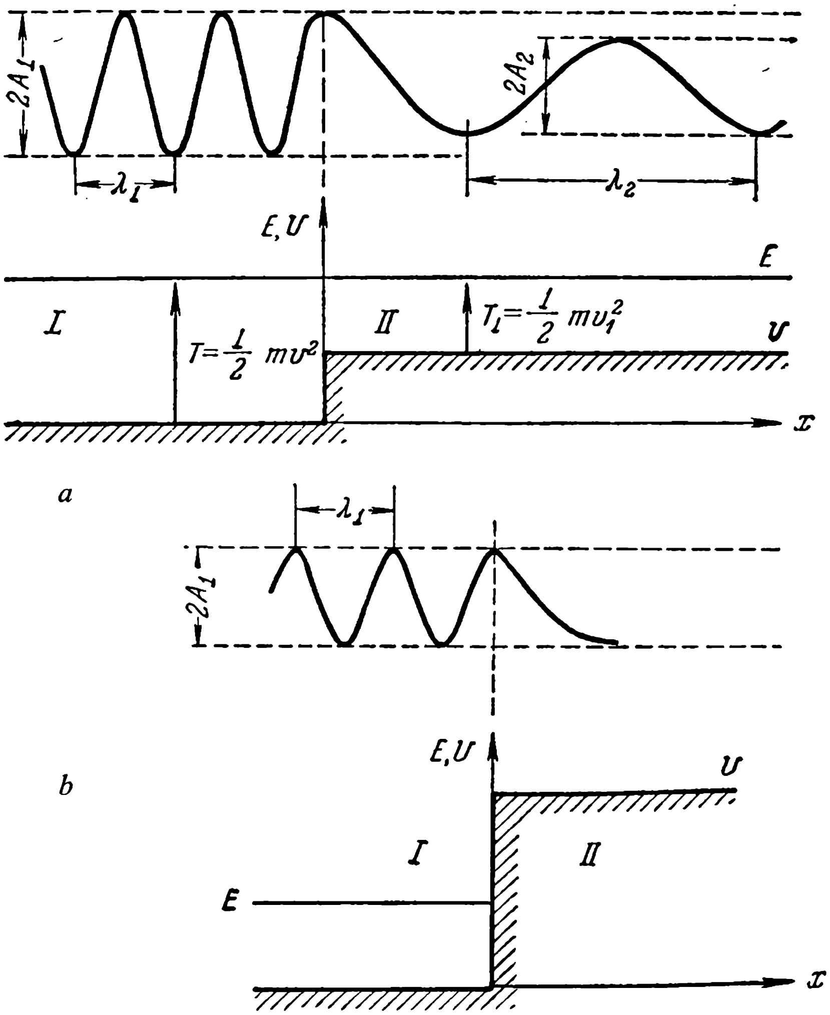


FIG. 2. (*a*) Passage of a particle across a low potential step; (*b*) reflection at the boundary of a high step.

$$\psi_1''(x)+k_1^2\psi_1(x) = 0; \quad \psi_2''(x)+k_2^2\psi_2(x) = 0, \qquad \ldots [1.58]$$

where ψ_1 and ψ_2 are the wave functions of the particle in fields I and II, and the parameters k_1 and k_2 are given by the expressions

$$k_1 = \frac{2\pi}{h}\sqrt{2m(E-U_1)}; \quad k_2 = \frac{2\pi}{h}\sqrt{2m(E-U_2)}. \qquad \ldots [1.59]$$

The general solutions of equations [*58*] have the form

$$\left.\begin{aligned} \psi_1(x) &= A_1 e^{ik_1x}+B_1 e^{-ik_1x} \ (x \leqq a) \\ \psi_2(x) &= A_2 e^{ik_2x}+B_2 e^{-ik_2x} \ (x \geqq a). \end{aligned}\right\} \qquad \ldots [1.60]$$

For arbitrary values of the coefficients A_i, B_i, the branches ψ_1 and ψ_2 do not, in general, form a single continuous curve $\psi(x)$ (Fig. 3*a*). For continuity (Fig. 3*b*) we require that the values of the wave function and its derivative at the interface between the fields should satisfy:

$$\psi_1(a) = \psi_2(a) \text{ and } \psi_1'(a) = \psi_2'(a). \qquad \ldots [1.61]$$

These impose limitations on the coefficients A_i, B_i, which must satisfy

$$\left.\begin{aligned} A_1 e^{ik_1 a} + B_1 e^{-ik_1 a} &= A_2 e^{ik_2 a} + B_2 e^{-ik_2 a} \\ k_1 A_1 e^{ik_1 a} - k_1 B_1 e^{-ik_1 a} &= k_2 A_2 e^{ik_2 a} - k_2 B_2 e^{-ik_2 a} \end{aligned}\right\} \qquad \ldots [1.62]$$

value E of the particle energy in comparison with the height U of the step. Mathematically, the nature of the solution of [*60*] is determined by whether the coefficients k_1 and k_2 are real or imaginary. If the particle energy E is less than the value of the smaller of the potentials, then energy considerations show that the particle cannot penetrate into the fields in question, since in that case its kinetic energy would be negative, k_1 and k_2 would be imaginary and the solutions of [*60*] would be sums of two exponential functions, such as $\psi(x) = Ae^{kx} + Be^{-kx}$, where k is real. Then the condition that the function be everywhere finite can only be satisfied if $A = B = 0$.

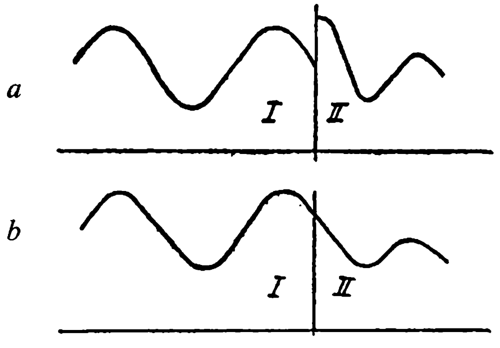


FIG. 3. 'Fitting together' of wave functions: (*a*) $\psi_1 + \psi_2$; (*b*) a continuous function ψ.

At the boundary of the step the de Broglie waves undergo *reflection* and *refraction* analogous to the reflection and refraction of light at a boundary separating two optically different media. Depending on the conditions of the problem, reflection may be partial or total; this is determined by the value of the reflection coefficient R, which is equal to the ratio between the intensities of the reflected and incident waves:

$$R = \frac{|B_1|^2}{|A_1|^2}. \qquad \ldots [1.63]$$

When the reflection is incomplete there is a probability that the particle may penetrate into the second medium; this is given by the transmission coefficient D of the boundary of separation. We shall find the expression giving the transmission coefficient for a stream of particles moving normally to the surface of separation and with a change of velocity at it. Let v_1 and v_2 be the velocities of the particles and ρ_1 and ρ_2 the densities of the streams in the two media. The number of particles crossing 1 cm.^{2} of the surface of separation per sec. is equal to the number of particles in cylinders of height equal to the velocities of the particles. Hence

$$D = \frac{n_2}{n_1} = \frac{\rho_2 v_2}{\rho_1 v_1}.$$

Since the densities of the streams are proportional to the squares of the amplitudes of the wave functions, the transmission coefficient for a de Broglie wave will be

$$D = \frac{|A_2|^2}{|A_1|^2}\left(\frac{v_2}{v_1}\right) = \frac{|A_2|^2}{|A_1|^2}\left(\frac{k_2}{k_1}\right). \qquad \ldots [1.64]$$

The total probability that a particle be either transmitted or reflected is unity, so

$$R+D = \frac{k_1\,|B_1|^2+k_2\,|A_2^2|}{k_1\,|A_1|^2} = 1. \qquad \ldots [1.65]$$

On crossing a low step a particle is slowed down and its kinetic energy diminished. The particle velocity and the de Broglie wavelength are inversely proportional (see [*50*] and [*51*]). Consequently the refractive index of de Broglie waves is

$$n = \frac{\lambda_1}{\lambda_2} = \frac{v_2}{v_1} = \frac{k_2}{k_1} = \sqrt{\frac{E-U_2}{E-U_1}}.$$

Setting $U_1 = 0$, we obtain

$$n = \sqrt{1-\frac{U}{E}}. \qquad \ldots [1.66]$$

The refractive index is determined by the ratio of the height of the step to the particle energy.

Refraction of de Broglie waves is observed during electron diffraction in crystals and is used to measure the internal potential of crystals (Ch. 8, § 1). The magnitude of this internal potential amounts to 15-20 V, so refraction can be readily observed only in diffraction of slow electrons with energies of the order of 100-300 eV.

Let us consider the reflection and transmission coefficients for a particle passing a low, and then a high, step.

1. Low step ($E>U$). The wave numbers of a particle in fields I and II are respectively

$$k_1 = \frac{2\pi}{h}\sqrt{2mE}; \quad k_2 = \frac{2\pi}{h}\sqrt{2m(E-U)}. \qquad \ldots [1.67]$$

The wave functions take the form

$$\psi_1 = A_1 e^{ik_1x}+B_1 e^{-ik_1x}; \quad \psi_2 = A_2 e^{ik_2x}. \qquad \ldots [1.68]$$

In field II only the transmitted wave is propagated, so $B_2 = 0$. The wave functions ψ_1 and ψ_2 represent travelling waves. The passage of a particle across a step is accompanied by a change in the wavelength, as shown in Fig. 2*a*. The continuity conditions for the function $\psi(x)$ and its derivative at the boundary of the jump give two equations for the coefficients:

$$A_1+B_1 = A_2; \quad k_1A_1-k_1B_1 = k_2A_2; \qquad \ldots [1.69]$$

a third equation is obtained from the normalisation condition. The elimination of A_2 from equation [*69*] enables us to find the ratio of the amplitudes B_2/A_1 and the reflection coefficient

$$R = \frac{|B_1|^2}{|A_1|^2} = \left(\frac{k_1-k_2}{k_1+k_2}\right)^2 = \left(\frac{1-\sqrt{1-q}}{1+\sqrt{1-q}}\right)^2, \text{ where } q = \frac{U}{E}. \qquad \ldots [1.70]$$

A calculation using this formula gives the following values for R:

$q = \frac{E}{U}$	1·00	1·11	1·25	2	10
R	1	0·2700	0·1459	0·0296	0·0007

and shows that even when the particle energy is much greater than the height of the step the reflection coefficient is appreciably different from zero. Thus, for $q = 2$, $R = 3\%$. A quantum particle differs in principle from a classical particle, which must without fail pass across a step if its energy exceeds the height of the step.

2. High step ($E < U$). One of the wave numbers in this case is imaginary:

$$k_1 = \frac{2\pi}{h}\sqrt{2mE}; \quad k_2 = ik; \text{ where } k = \frac{2\pi}{h}\sqrt{2m(U-E)}. \qquad \ldots [1.71]$$

The wave functions take the form

$$\psi_1 = A_1 e^{ik_1x} + B_1 e^{-ik_1x}; \quad \psi_2 = A_2 e^{-kx} + B_2 e^{kx} = A_2 e^{-kx}. \qquad \ldots [1.72]$$

In field I the wave function represents the superposition of an incident and a reflected wave, and in field II the sum of two exponential functions. From the condition that ψ remain finite it follows that $B_2 = 0$. The continuity conditions for $\psi(x)$ and its derivative lead to equations of the type [*69*]. The reflection coefficient is found from [*70*] on putting $k_2 = ik$:

$$R = \frac{|k_1 - ik|^2}{|k_1 + ik|}. \qquad \ldots [1.73]$$

R being a complex quantity, we take its modulus $|R|$:

$$|R|^2 = RR^* = 1, \text{ whence } D = 0. \qquad \ldots [1.74]$$

Formula [*74*] shows that at the boundary of the step total reflection is observed, and the particle cannot move into field II.

However, in distinction to the classical case, the particle does penetrate into a narrow border region of field II, the probability of penetration being proportional to the square of the modulus of the corresponding wave function:

$$w(x) = |\psi_2|^2 = A_2^2 e^{-2kx} = A_2^2 e^{-\frac{4\pi}{h}\sqrt{2m(U-E)}\,x}. \qquad \ldots [1.75]$$

A calculation using this formula, with an excess of step over particle energy of $U - E = 1$ eV, gives the following values for the probability:

x	1	5	10 Å
$w(x)$	0·29	0·005	$4{\cdot}5 \times 10^{-8}$.

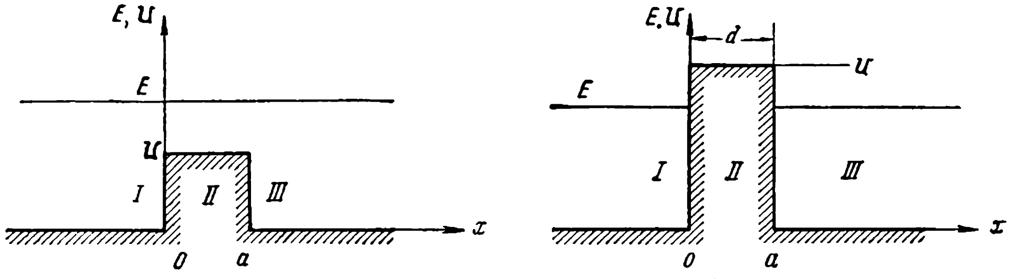
The exponential function [*75*] produces a sharp drop in w with increasing depth of penetration. However, within a narrow boundary layer, the probability of the particle penetrating has a reasonably large value. This is analogous to the penetration of an electromagnetic field into a second medium during total internal reflection of light.

4. Passage of a Particle through a Potential Barrier

Let us consider the passage of a particle through a potential barrier (Fig. 4) consisting of a potential step of finite width a. The barrier divides space into three regions:

$$\begin{array}{lll} \text{I.} & x \leqq 0; & U_1 = 0; \\ \text{II.} & 0 \leqq x \leqq a; & U_2 = U; \\ \text{III.} & a \leqq x; & U_3 = 0. \end{array}$$

Depending on the ratio of the particle energy to the barrier height, we must distinguish: (*a*) low ($E > U$) and (*b*) high ($E < U$) barriers. A high barrier of sufficient width is as opaque as a potential step. However a narrow barrier can be penetrated if its width is comparable with the dimensions of the region within which the wave function is appreciably non-zero. The transmission coefficient D increases rapidly as barrier width and height are diminished.



a *b*

FIG. 4. Passage of a particle through a potential barrier: (*a*) low, (*b*) high.

1. Low Barrier ($E > U$). A low barrier (Fig. 4*a*) may be looked upon as a region of perturbation, in passing across which a particle is slowed down at first, and then regains its original velocity. We shall find the reflection (scattering) coefficient for the particle. The wave numbers for the particle are respectively:
for regions I and III:

$$k_1 = \frac{2\pi}{h}\sqrt{2mE};$$

for region II:

$$k_2 = \frac{2\pi}{h}\sqrt{2m(E-U)}. \qquad \ldots [1.76]$$

The wave functions represent travelling waves; in region III there is only a forward wave:

$$\psi_1 = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}; \quad \psi_2 = A_2 e^{ik_2 x} + B_2 e^{-ik_2 x}; \quad \psi_3 = A_3 e^{ik_1 x}. \qquad \ldots [1.77]$$

The continuity conditions for ψ and its derivative give four relationships between the amplitudes:

$$\left.\begin{array}{l} A_1 + B_1 = A_2 + B_2; \quad k_1 A_1 - k_1 B_1 = k_2 A_2 - k_2 B_2 \\ A_2 e^{ik_2 a} + B_2 e^{-ik_2 a} = A_3 e^{ik_1 x}; \\ k_2 A_2 e^{ik_2 a} - k_2 B_2 e^{-ik_2 a} = k_1 A_3 e^{ik_1 a} \end{array}\right\} \qquad \ldots [1.78]$$

a fifth relationship can be obtained from the normalisation condition. Eliminating A_3 from the last two equations, we obtain

$$\frac{B_2}{A_2} = \frac{k_2 - k_1}{k_2 + k_1} e^{2ik_2 a}. \qquad \ldots [1.79]$$

From the first two equations we find

$$\frac{B_2}{A_2} = \frac{(k_2-k_1)A_1+(k_2+k_1)B_1}{(k_2+k_1)A_1+(k_2-k_1)B_1}. \qquad \ldots [1.80]$$

From [79] and [80] we have

$$\frac{B_1}{A_1} = \frac{i(k_2^2-k_1^2)}{i(k_1^2+k_2^2)+2k_1k_2 \cot k_2 a}. \qquad \ldots [1.81]$$

The scattering coefficient is

$$R = \frac{|B_1|^2}{|A_1|^2} = \frac{B_1 B_1^*}{A_1 A_1^*} = \frac{(k_1^2-k_2^2)^2}{(k_1^2+k_2^2)^2+4k_1^2k_2^2 \cot^2 k_2 a}. \qquad \ldots [1.82]$$

The scattering coefficient takes especially high values when

$$k_2 a = n\pi \quad (n = 1, 2, 3 \ldots), \qquad \ldots [1.83]$$

which corresponds to certain definite values of the particle energy. In this case the de Broglie wavelength satisfies the relationship

$$n\lambda = 2a, \qquad \ldots [1.84]$$

which shows that within the width of the barrier there must be an integral number of half wavelengths.

2. High Barrier ($E < U$). In this case (Fig. 4*b*) the wave numbers have the following values:

for regions I, III

$$k_1 = \frac{2\pi}{h}\sqrt{2mE}; \qquad \ldots [1.85]$$

for region II

$$k_2 = ik,$$

where

$$k = \frac{2\pi}{h}\sqrt{2m(U-E)}.$$

In the first two regions the wave functions represent travelling waves. In region II k_2 is imaginary and here the wave function is the sum of two exponential functions:

$$\psi_1 = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}; \quad \psi_2 = A_2 e^{-kx} + B_2 e^{kx}; \quad \psi_3 = A_3 e^{ik_1 x}. \qquad \ldots [1.86]$$

The continuity conditions give

$$\left.\begin{aligned} &A_1 + B_1 = A_2 + B_2; \quad ik_1 A_1 - ik_1 B_1 = -kA_2 + kB_2; \\ &A_2 e^{-ka} + B_2 e^{ka} = A_3 e^{ik_1 a}; \\ &-kA_2 e^{-ka} + kB_2 e^{ka} = ik_1 A_3 e^{ik_1 a}. \end{aligned}\right\} \qquad \ldots [1.87]$$

Solving this system enables us to find the ratio of the amplitudes:

$$\frac{A_3}{A_1} = \frac{4ik_1ke^{ik_1a}}{(k_1+ik)^2e^{ka}-(k_1-ik)e^{-ka}}. \qquad \ldots [1.88]$$

Introducing hyperbolic functions:

$$\sinh x = \tfrac{1}{2}(e^x - e^{-x}), \quad \cosh x = \tfrac{1}{2}(e^x + e^{-x}), \qquad \ldots [1.89]$$

we obtain

$$\frac{A_3}{A_1} = \frac{2ik_1ke^{ik_1a}}{(k_1^2-k^2)\sinh^2 ka + 2ik_1k\cosh ka}. \qquad \ldots [1.90]$$

Multiplying by the complex conjugate quantity, we obtain the transmission coefficient of the barrier:

$$D = \frac{A_3A_3^*}{A_1^2} = \frac{4k_1^2k^2}{(k_1^2-k^2)^2\sinh^2 ka + 4k_1^2k^2\cosh^2 ka}. \qquad \ldots [1.91]$$

Since $\cosh^2 x - \sinh^2 x = 1$,

$$D = \frac{4k_1^2k^2}{(k_1^2+k^2)^2\sinh^2 ka + 4k_1^2k^2}. \qquad \ldots [1.92]$$

For sufficiently large x, as can be seen from [*89*], $\sinh x \approx \frac{1}{2}e^x$; and using in addition the approximation $k_1 \sim k$, we obtain

$$D \approx \frac{4}{e^{2ka}+4}.$$

Neglecting the second term in the denominator, we find for the transmission coefficient:

$$D \approx e^{-2ka} = Ae^{-\frac{4\pi}{h}\sqrt{2m(U-E)}\,a}, \qquad \ldots [1.93]$$

which is the same as the expression [*75*] which gave the probability of a particle penetrating into the region beyond a high step. Formula [*93*] shows that a particle can pass across a barrier even when its energy is less than the height of the barrier. The energy of a particle after crossing a barrier is equal to its original energy (Fig. 5). This passage through a potential barrier is sometimes called *the tunnel effect.* The penetrability of a barrier depends to a great extent on its height and especially on its width. Thus for an energy difference $U-E = 5$ eV the transmission varies with width in the following manner:

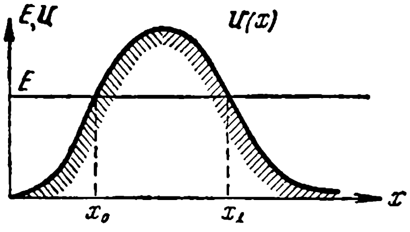


FIG. 5. The tunnel effect. Passage of a particle through a barrier of arbitrary shape.

a	1·0	1·5	2·0	5·0 Å
D	0·10	0·03	0·008	$5{\cdot}5\times10^{-7}$.

Thus where the width of the barrier is of the order of atomic radii, 1-2 Å, the penetrability is appreciable. Passage through a potential barrier is essential for an understanding of such phenomena as contact potential difference and the extraction of electrons from a metal by an electric field. The α-decay of heavy nuclei is explained by leakage of α-particles through the potential barrier of the nucleus. The transparency of a barrier of arbitrary shape is given by the formula:

$$D = Ae^{-\frac{4\pi}{h}\sqrt{2m}\int\limits_{x_1}^{x_2}\sqrt{[U(x)-E]}dx} \qquad \ldots [1.94]$$

which is a generalisation of formula [*93*].

5. Particle in a Rectangular Potential Well

Formulation of the Problem. Weak and Strong Binding

Let us consider a potential well (Fig. 6) with cross-section $U = U(x)$, maximum depth U_0 and width a. We shall regard the horizontal sections of the potential curve outside the well as extending to infinity. Depending on the relationship between the particle energy E and the depth of the well, we must distinguish two cases: the particle energy either: (1) exceeds the well depth ($E > U_0$), or (2) is less than it ($E < U_0$). The cases correspond to the weak and strong binding respectively.

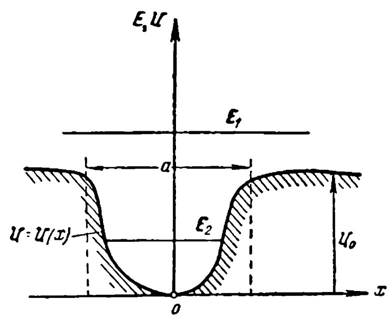


FIG. 6. Particle in a potential well.

With weak binding, the potential well can be looked on as a negative barrier, across which the particle passes with energy greater than its height. The particle then moves like a free particle everywhere except in a region of potential perturbation where its motion is at first accelerated, then retarded. Variations in the kinetic energy and velocity of the particle are accompanied by variations in the de Broglie wave length and in the amplitude of the wave function. After passage across the region of perturbation, the original velocity of motion is restored. The perturbation effect increases with diminution of the total particle energy. If the particle energy becomes less than the well depth, then the particle is trapped by the well and remains firmly bound to it (strong binding).

With strong binding the particle has its motion fundamentally altered in character: its motion in space is limited to the region of the potential well. The range of variation of the coordinates of the particle is determined by the dimensions of the well. Reflection from the walls of the well leads to the particle's motion being *periodic* in time, thus imposing *quantisation* conditions on the particle energy and momentum. The problem of the motion of a particle in a potential well has many applications in various branches of solid state, atomic and nuclear physics.

Strong Binding. One-dimensional Potential Box of Infinite Depth

The simplest well is one with plane sides, separating fields of constant potential (potential box). Fig. 7 shows the cross-section of a potential box of width a with walls of infinite height. A particle in such a box will be perpetually bound to it. The potential box divides space into two regions: region I—the space inside the box, and region II—the rest of space outside the box. In region I $0 \leqq x \leqq a$, $U_1 = 0$; in region II $x \leqq 0$ or $x \geqq a$, $U_2 = \infty$. The wave numbers are given by formulae [*71*]. When $U = U_2 = \infty$ we have $k = \infty$. The solutions are the wave functions [*72*]. When $k = \infty$, $\psi_2 = 0$ everywhere, in this case on the walls of the box, $\psi_2(0) = \psi_2(a) = 0$.

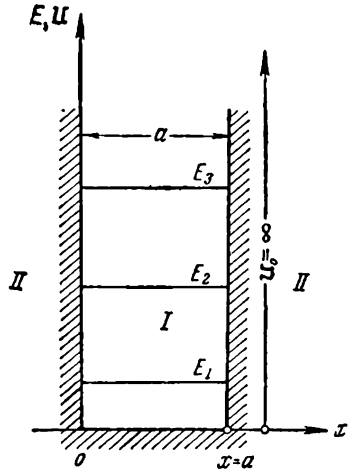


FIG. 7. Potential box with infinitely high walls.

From the continuity of the wave function it follows that

$$\psi_1(0) = \psi_1(a) = 0.$$

From these boundary conditions we obtain the equations for the coefficients:

$$\left.\begin{aligned} A_1 + B_1 &= 0 \\ A_1 e^{ik_1a} + B_1 e^{-ik_1a} &= 0 \end{aligned}\right\}. \qquad \ldots [1.95]$$

The system of homogeneous equations [*95*] is compatible provided the determinant formed from the coefficients in the equations vanishes:

$$\begin{vmatrix} 1 & 1 \\ e^{ik_1a} & e^{-ik_1a} \end{vmatrix} = 0, \qquad \ldots [1.96]$$

which leads to an equation for k_1 and, consequently, for the particle energy

$$e^{2ik_1a} = 1. \qquad \ldots [1.97]$$

The roots of this equation satisfy the relationship

$$k_1 a = n\pi. \qquad \ldots [1.98]$$

Substituting in this the value of k_1 from [*72*], we find that equations [*95*] are solvable if

$$E_n = \frac{n^2 h^2}{8ma^2}, \; (n = 1, 2, 3, \ldots), \qquad \ldots [1.99]$$

showing that inside the potential box the particle energy is *quantised*. The energy spectrum of the particle is discrete (Fig. 7). The distances between levels increase with increasing quantum numbers and are proportional to the differences of their squares:

$$E_1 - E_2 = \frac{h^2}{8ma^2}(n_1^2 - n_2^2).$$

If the dimensions of the potential box are considerable and amount to macroscopic magnitudes, for example $a = 1$ cm, then even for the lightest particles—electrons—we obtain

$$E_n = 5{\cdot}4 \times 10^{-27}\; n^2 \text{ erg} = 3{\cdot}37 \times 10^{-15}\; n^2 \text{ eV}.$$

Hence the distances between neighbouring levels are

$$\Delta E_n \approx 6{\cdot}7 \times 10^{-15}\, n \text{ eV}.$$

The levels are very close to one another and form a quasi-continuous band. The discrete nature of the spectrum of a microparticle in a box is apparent only in the case when the dimensions of the box are comparable with atomic dimensions. For a box with $a = 5$ Å, we obtain

$$E_n = 0{\cdot}68n^2 \text{ eV},$$

whence

$$\Delta E_n = 1{\cdot}36n \text{ eV}.$$

We obtain the wave functions for a particle in a box from [72] by calculating the relationship between the coefficients, [*95*]:

$$|\psi_n(x)| = 2A_1 \sin\frac{n\pi x}{a}. \qquad \ldots [\mathit{1.100}]$$

The wave functions of the particle represent standing waves. It can easily be proved that the functions ψ_n are all mutually orthogonal.

From the normalisation condition

$$\int_{-\infty}^{+\infty} |\psi_n|^2 dx = 1 \text{ we obtain } 2A_1 = \sqrt{\frac{2}{a}}.$$

Consequently the normalised functions will be

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\frac{\pi n x}{a}. \qquad \ldots [\mathit{1.101}]$$

The probability density distribution for a particle in a box is

$$w_n(x) = |\psi_n(x)|^2 = \frac{2}{a}\sin^2\frac{\pi n x}{a}. \qquad \ldots [\mathit{1.102}]$$

The corresponding curves w_n are shown in Fig. 8.

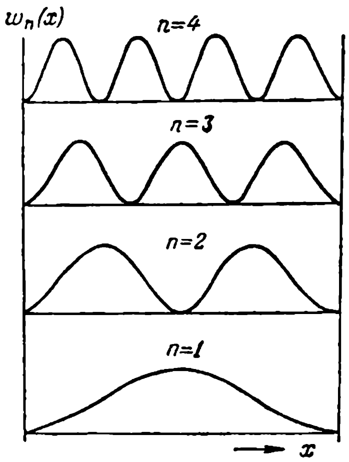


FIG. 8. Probability density distribution for a particle in a potential box.

Generalisation to a Three-dimensional Box

Let us suppose that the potential box has the shape of a rectangular parallelepiped with sides a, b, c. The potential energy of the particle is $U_1 = 0$ inside the box and $U_2 = \infty$ outside it. Schrödinger's equation for the three-dimensional case is

$$\nabla^2\psi(x,y,z) + \frac{8\pi^2 m}{h}(E-U)\psi(x,y,z) = 0. \qquad \ldots [\mathit{1.103}]$$

The components of motion of a free particle along the X, Y, and Z axes are independent. So we shall seek a solution of equation [*103*] in the form of a function with separated variables:

$$\psi(x,y,z) = X(x)\,Y(y)\,Z(z), \qquad \ldots [\mathit{1.104}]$$

which corresponds to the expression for the probability of a compound event. Substituting [*104*] into [*103*] and dividing both sides of the equation obtained by $\psi = XYZ$, we then have

$$\left(\frac{1}{X}\frac{d^2X}{dx^2}-\frac{8\pi^2m}{h^2}U_x\right)+\left(\frac{1}{Y}\frac{d^2Y}{dy^2}-\frac{8\pi^2m}{h^2}U_y\right)+\left(\frac{1}{Z}\frac{d^2Z}{dz^2}-\frac{8\pi^2m}{h^2}U_z\right) = -\frac{8\pi^2m}{h^2}E. \qquad \ldots [1.105]$$

Starting from the equivalence of the three coordinate axes, we have here separated the magnitude of the potential U into three components U_x, U_y, U_z, which may be chosen arbitrarily, provided the relationship

$$U_x+U_y+U_z = U \qquad \ldots [1.106]$$

is maintained.

The L.H.S. of [*105*] has broken down into the sum of three terms, each of which separately depends only on a single variable. The R.H.S. of [*105*] is constant. The equality expressed by [*105*] is possible only if each of the terms on the L.H.S. is a constant. Thus equation [*105*] breaks down into three independent equations, which we shall write in the form:

$$\left.\begin{aligned}\frac{1}{X}\frac{d^2X}{dx^2}-\frac{8\pi^2m}{h^2}U_x &= -\frac{8\pi^2m}{h^2}E_x\\ \frac{1}{Y}\frac{d^2Y}{dy^2}-\frac{8\pi^2m}{h^2}U_y &= -\frac{8\pi^2m}{h^2}E_y\\ \frac{1}{Z}\frac{d^2Z}{dz^2}-\frac{8\pi^2m}{h^2}U_z &= -\frac{8\pi^2m}{h^2}E_z\end{aligned}\right\}. \qquad \ldots [1.107]$$

Equation [*103*] for a three-dimensional box has thus broken down into three equations, corresponding to three one-dimensional boxes. Using the solution found earlier (see [*99*] and [*101*]), we obtain for the present case:

$$E_{nx} = \frac{n_x^2h^2}{8ma^2};\quad E_{ny} = \frac{n_y^2h^2}{8mb^2};\quad E_{nz} = \frac{n_z^2h^2}{8mc^2}, \qquad \ldots [1.108]$$

$$X_n = \sqrt{\frac{2}{a}}\sin\frac{n_x\pi x}{a};\quad Y_n = \sqrt{\frac{2}{b}}\sin\frac{n_y\pi y}{b};\quad Z_n = \sqrt{\frac{2}{c}}\frac{n_z\pi z}{c}. \qquad \ldots [1.109]$$

where $n_x, n_y, n_z = 1, 2, 3 \ldots$

The resulting wave function for the particle follows from [*104*] and [*108*]:

$$\psi_{n_xn_yn_z}(x,y,z) = \sqrt{\frac{8}{v}}\sin\frac{n_x\pi x}{a}\sin\frac{n_y\pi y}{b}\sin\frac{n_z\pi z}{c},\quad (v = abc). \qquad \ldots [1.110]$$

From [*106*] and [*108*] we obtain the expression for the particle energy:

$$E_{n_x,n_y,n_z} = \frac{h^2}{8m}\left(\frac{n_x^2}{a^2}+\frac{n_y^2}{b^2}+\frac{n_z^2}{c^2}\right). \qquad \ldots [1.111]$$

The wave function and the particle energy in the three-dimensional potential box are determined by three quantum numbers (n_x, n_y, n_z), which correspond to the three degrees of freedom possessed by the particle. The wave function has the form of a three-dimensional standing wave. The particle energy forms a discrete spectrum. In order to give a graphic representation of the spectrum, we shall construct a three-dimensional model having the form of a space lattice (Fig. 9), with periods a^*, b^*, c^* which are the reciprocals of the lengths of the sides of the potential box:

$$a^* = \frac{1}{a}, \quad b^* = \frac{1}{b}, \quad c^* = \frac{1}{c}. \qquad \ldots [1.112]$$

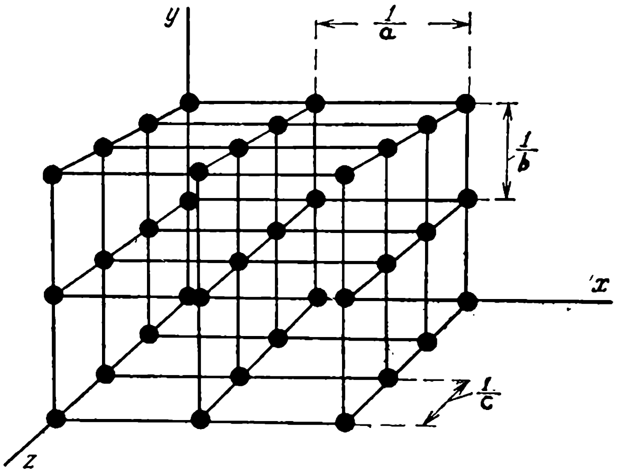


FIG. 9. Representation of the energy spectrum of a particle in a three-dimensional potential box.

Each node in this lattice is specified by three integral coordinates n_x, n_y, n_z and corresponds to one of the possible quantised states of the particle. The square of the length of the radius vector r in this lattice,

$$r^2 = n_x^2 a^{*2} + n_y^2 b^{*2} + n_z^2 c^{*2} = \frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \qquad \ldots [1.113]$$

is proportional to the particle energy:

$$E_{n_x, n_y, n_z} \equiv \frac{h^2}{8m} r^2_{n_x, n_y, n_z}. \qquad \ldots [1.114]$$

Potential Box of Finite Height

For a potential box of finite height there can be either weak or strong binding, depending on the particle energy (Fig. 10).

1. Weak Binding ($E > U$). The wave numbers of the particle have the following values:

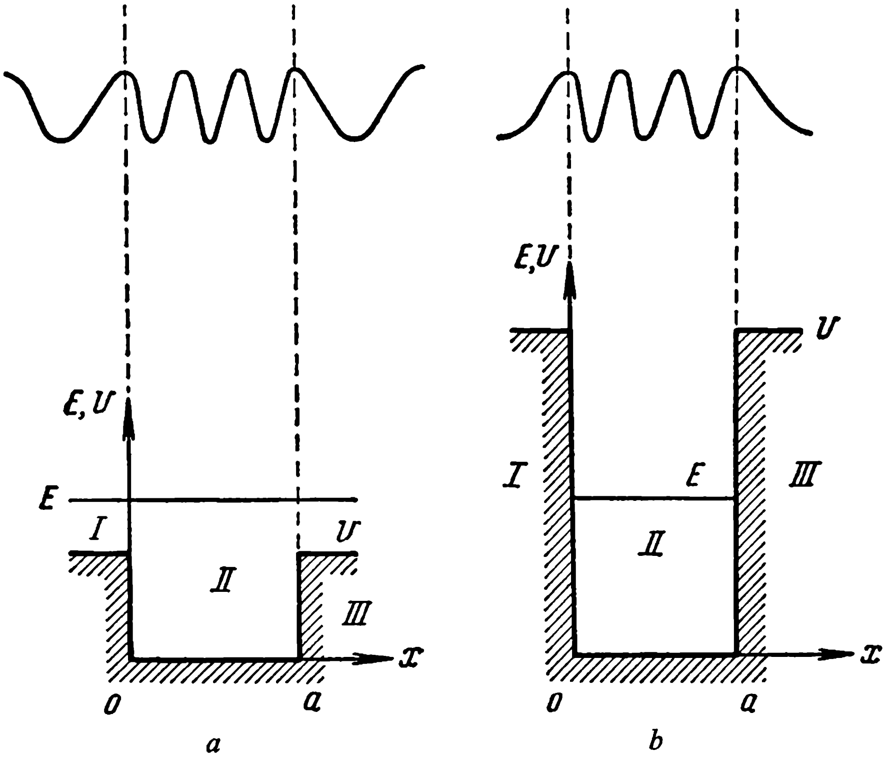
for regions I and III

$$x \leqq 0 \quad \text{and} \quad x \geqq a, \quad k_1 = \frac{2\pi}{h}\sqrt{2m(E-U)};$$

for region II

$$0 \leqq x \leqq a, \quad k_2 = \frac{2\pi}{h}\sqrt{2mE}. \qquad \ldots [1.115]$$

Apart from the transposition of the indices, these are the same as the wave numbers [*76*] for a low potential barrier, so the solutions of the problems of a weakly bound potential well, and of a low barrier, are analogous. The wave functions take the form of travelling waves as in [*77*], and the continuity conditions lead to the same equations as in [*78*]. Solving these equations enables us to find the scattering coefficient R. As can be seen from formula [*82*], R does not alter when the indices 1 and 2 are transposed. So positive and negative perturbations of the potential lead to the same value for the scattering coefficient.



a b

FIG. 10. Potential box of finite height: (*a*) weak binding, (*b*) strong binding.

2. *Strong Binding* ($E<U$). The wave numbers have the following values:

$$\left.\begin{array}{ll}\text{for regions I and III} & \\ & x\leqq 0 \text{ and } x\geqq a,\quad k_1 = ik, \\ & \text{where } k = \dfrac{2\pi}{h}\sqrt{2m(U-E)}; \\ \text{while for region II} & 0\leqq x\leqq a,\quad k_2 = \dfrac{2\pi}{h}\sqrt{2mE}.\end{array}\right\} \qquad \ldots [1.116]$$

In contradistinction to the case of a high barrier, the wave function for a deep potential well is represented by travelling waves only in region II, that is, inside the well, while in regions I and III the wave function is represented by the exponential functions

$$\psi_1 = A_1 e^{kx};\quad \psi_2 = A_2 e^{ik_2 x} + B_2 e^{-ik_2 x};\quad \psi_3 = B_3 e^{-kx}. \qquad \ldots [1.117]$$

The finiteness condition for the wave function imposes two limitations on the values of the amplitudes:

$$B_1 = 0 \quad \text{and} \quad A_3 = 0. \qquad \ldots [1.118]$$

The continuity conditions for ψ and its derivative yield the four following additional relationships between the amplitudes:

$$\begin{aligned} &A_1 = A_2+B_2; \quad kA_1 = ik_2A_2-ik_2B_2; \\ &A_2e^{ik_2a}+B_2e^{-ik_2a} = B_3e^{-ka}; \quad ik_2A_2e^{ik_2a}-ik_2B_2e^{-ik_2a} = -kB_3e^{-ka}. \end{aligned} \qquad \ldots [1.119]$$

Equations [*119*] represent a system of four homogeneous equations in four unknowns A_1, A_2, B_2 and B_3. This system is soluble if the determinant formed from the coefficients of the equations vanishes:

$$\begin{vmatrix} 1 & -1 & -1 & 0 \\ k & -ik_2 & ik_2 & 0 \\ 0 & e^{ik_2a} & e^{-ik_2a} & -e^{-ka} \\ 0 & ik_2e^{ik_2a} & -ik_2e^{-ik_2a} & ke^{-ka} \end{vmatrix} = 0. \qquad \ldots [1.120]$$

Formula [*120*], like formula [*96*] for the infinitely deep box, imposes limitations on the particle energy and contains within itself the quantisation conditions. Multiplying the first row by k and subtracting it from the second, we lower the rank of the determinant by one:

$$\begin{vmatrix} k-ik_2 & k+ik_2 & 0 \\ e^{ik_2a} & e^{-ik_2a} & -e^{-ka} \\ ik_2e^{ik_2a} & -ik_2e^{-ik_2a} & ke^{-ka} \end{vmatrix} = 0. \qquad \ldots [1.121]$$

Multiplying the second row by k and adding to it the third, we again lower the rank of the determinant:

$$\begin{vmatrix} k-ik_2 & k+ik_2 \\ (k+ik_2)e^{ik_2a} & (k-ik_2)e^{-ik_2a} \end{vmatrix} = 0. \qquad \ldots [1.122]$$

Expanding this determinant, we obtain after some reduction an equation of the form

$$\tan k_2a = -\frac{2kk_2}{k^2-k_2^2}. \qquad \ldots [1.123]$$

For an infinitely deep box $k = \infty$, and equation [*123*] simplifies to $\tan k_2a = 0$. The solution in this case will be $k_2a = n\pi$, which, allowing for the transposition of indices, is the same as [*98*].

Putting the values of k_1 and k_2 from [*116*] into [*123*], we obtain the equation for the particle energy:

$$\tan\left(\frac{2\pi}{h}\sqrt{2mEa}\right) = -\frac{2\sqrt{E(U-E)}}{U-2E}. \qquad \ldots [1.124]$$

This can be written in the form

$$\tan(\pi C\sqrt{x}) = -\frac{2\sqrt{x(1-x)}}{1-2x}, \quad (0 \leqq x \leqq 1), \qquad \ldots [1.125]$$

where

$$x = \frac{E}{U} \quad \text{and} \quad C = \frac{2a}{h}\sqrt{2mU}. \qquad \ldots [1.126]$$

In order to analyse equation [*125*] we shall represent it in graphical form (Fig. 11). The curves $P(x)$ represent the R.H.S. of the equation, while its L.H.S. for $C \geqq 1$ is a family of tangent curves whose intersections with the curve $P(x)$ determine the possible values of x, and give the sequence of quantised energy values $E_1, E_2, \ldots, E_n, \ldots$ for the particle. Equation [*125*] enables us to establish the particle's energy spectrum as a function of the parameters of the potential well; its depth, width, and also the mass of the particle. Variation of these parameters enables us to change the 'strength' of the bond.

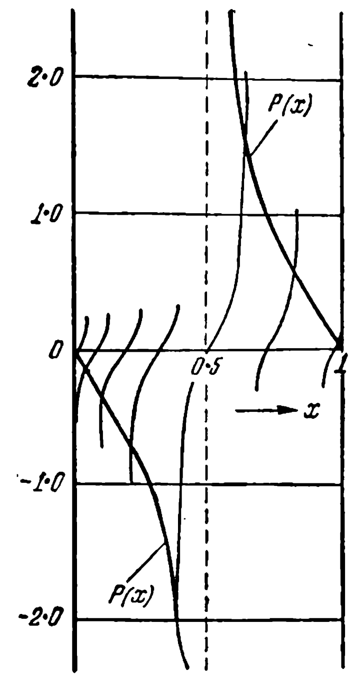


FIG. 11. Graphical solution of equation *1.125*.

6. The Harmonic Oscillator

The Classical Oscillator

Atoms in molecules and crystals are in a bound state. When the atoms are displaced from their equilibrium positions forces act on them which tend to restore them to these positions. For small displacements r the restoring force F may be regarded as proportional to the displacement:

$$F(r) = -fr, \qquad \ldots [1.127]$$

where f is the coefficient of the restoring force. The potential energy of the quasi-elastic oscillator,

$$U = \tfrac{1}{2}fr^2, \qquad \ldots [1.128]$$

varies parabolically (Fig. 12). The present problem is that of the motion of a particle in a parabolic potential well. In this case the classical particle performs harmonic oscillations:

$$r = A\cos 2\pi\nu t, \qquad \ldots [1.129]$$

where A is the amplitude, and ν the frequency of vibration,

$$\nu = \frac{1}{2\pi}\sqrt{\frac{f}{m}}. \qquad \ldots [1.130]$$

The kinetic and potential energies of the oscillator vary according to the laws:

$$\begin{aligned} T &= \tfrac{1}{2}mv^2 = \tfrac{1}{2}fA^2\sin^2 2\pi\nu t, \\ U &= \tfrac{1}{2}fr^2 = \tfrac{1}{2}fA^2\cos^2 2\pi\nu t. \end{aligned} \qquad \ldots [1.131]$$

The mean values of these energies over the period of a single oscillation are equal: $T = U$. The total energy of the oscillator,

$$E = T+U = \tfrac{1}{2}fA^2, \qquad \ldots [1.132]$$

is proportional to the square of the amplitude of oscillation; its values are not subject to any special limitations. The classical oscillator has a continuous energy spectrum.

The Quantum Oscillator. Hermite's Equation

Schrödinger's equation for the harmonic oscillator has the form

$$\frac{d^2\psi}{dr^2}+\frac{8\pi^2 m}{h^2}\left(E-\tfrac{1}{2}fr^2\right)\psi = 0. \qquad \ldots [1.133]$$

We write

$$\alpha = \frac{8\pi^2 mE}{h^2}, \quad \beta = \frac{2\pi\sqrt{mf}}{h} \qquad \ldots [1.134]$$

and introduce the new variable

$$\xi = \sqrt{\beta}r. \qquad \ldots [1.135]$$

Equation [*133*] now takes the form

$$\frac{d^2\psi}{d\xi^2}+\left(\frac{\alpha}{\beta}-\xi^2\right)\psi = 0. \qquad \ldots [1.136]$$

For large displacements, when α/β can be ignored in comparison with ξ^2, equation [*136*] simplifies to

$$\frac{d^2\psi}{d\xi^2} = \xi^2\psi. \qquad \ldots [1.137]$$

We considered earlier the simpler equation [*45*] with a constant coefficient, giving a wave function whose solutions are exponential. By analogy, we shall take as an approximate solution of equation [*137*] the function:

$$\psi(\xi) = Ce^{\pm\frac{\xi^2}{2}}, \qquad \ldots [1.138]$$

where C is a constant. The second derivative of this function,

$$\frac{d^2\psi}{d\xi^2} = C(\xi^2\pm 1)e^{\pm\frac{\xi^2}{2}} \qquad \ldots [1.139]$$

satisfies equation [*137*] in the case when unity can be neglected in comparison with ξ^2. Thus, [*138*] can be looked on as an asymptotic solution of [*137*] for large displacements. We shall try to improve this solution by replacing the constant coefficient C in [*138*] by a function:

$$\psi(\xi) = U(\xi)e^{-\frac{1}{2}\xi^2}. \qquad \ldots [1.140]$$

Substitution of [*140*] into [*136*] shows that the function [*140*] is a solution of the initial equation [*136*] provided $U(\xi)$ in its turn satisfies the following equation:

$$\frac{d^2U}{d\xi^2}-2\xi\frac{dU}{d\xi}+\left(\frac{\alpha}{\beta}-1\right)U = 0. \qquad \ldots [1.141]$$

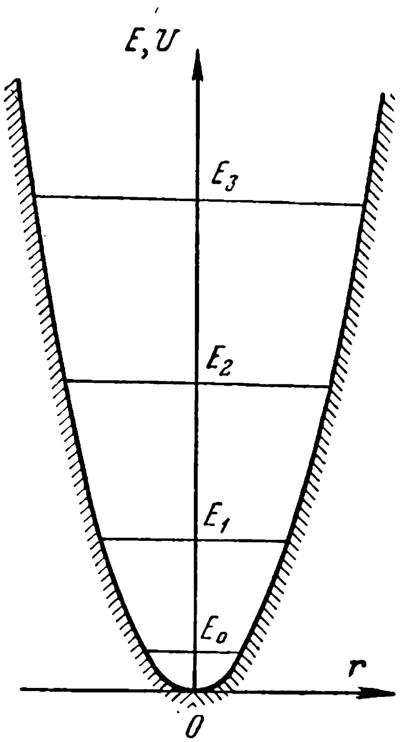


FIG. 12. Energy spectrum of a quantised harmonic oscillator.

We introduce the notation

$$\frac{\alpha}{\beta}-1 = 2n, \qquad \ldots [1.142]$$

by means of which equation [*141*] can be written in the form

$$U''-2\xi U'+2nU = 0. \qquad \ldots [1.143]$$

Equation [*143*] is Hermite's equation, which can have solutions only for integral values of the parameter n:

$$n = 0, 1, 2, \ldots \qquad \ldots [1.144]$$

Since α involves the energy E, formula [*142*] contains the quantisation condition for the total energy of the oscillator. Substituting into [*142*] the values of α and β from [*134*] and the value for the frequency of oscillation from [*130*], we obtain

$$E_n = (n+\tfrac{1}{2})\, h\nu. \qquad \ldots [1.145]$$

Formula [*145*] shows that the energy of a quantum oscillator varies discontinuously. The energy spectrum is equally spaced (Fig. 12). The distances between neighbouring levels are the same and are equal to the energy of a single quantum:

$$\varepsilon = h\nu. \qquad \ldots [1.146]$$

Formula [*146*] expresses the principle contained in Planck's quantum postulate. The novelty in formula [*145*] is the half quantum number, which determines the ground-state energy of the oscillator:

$$E_0 = \tfrac{1}{2}h\nu. \qquad \ldots [1.147]$$

This energy is conserved even at the absolute zero of temperature.

The existence in the oscillator of 'zero point' energy and of the corresponding amplitude of the residual 'zero point' oscillations is a consequence of the *principle of indeterminacy*. If at absolute zero the vibrations of the oscillator ceased completely, then it would be possible to make a simultaneous accurate determination both of the coordinate ($x = 0$) and of the momentum ($p = 0$). The existence of 'zero point' energy in crystals, which amount to a system of oscillators, is confirmed by measurements of the scattering intensity of X-rays in crystals at low temperatures, and also by the variation with temperature of the thermal

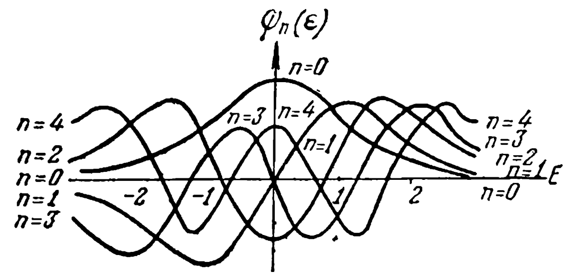


FIG. 13. The first five eigenfunctions of the oscillator.

capacity of molecules and crystals, in particular the vanishing of the thermal capacity of solids at absolute zero.

The normalised eigenfunctions for the harmonic oscillator are expressed in terms of Hermite polynomials:

$$\psi_n(\xi) = \left(\frac{1}{2^n n!}\sqrt{\frac{\beta}{\pi}}\right)^{\frac{1}{2}} H_n(\xi) e^{\frac{-\xi^2}{2}}. \qquad \ldots [1.148]$$

The first five eigenfunctions of the oscillator are shown in Fig. 13.

The mean displacement of the particle from its equilibrium position is

$$\bar{x} = \frac{1}{\sqrt{\beta}}\bar{\xi} = \frac{1}{2^n n! \sqrt{\pi\beta}} \int\limits_{-\infty}^{+\infty} H_n(\xi)\xi H_n(\xi) e^{-\xi^2} d\xi. \qquad \ldots [1.149]$$

For Hermite polynomials there is a recurrence formula:

$$\xi H_n(\xi) = nH_{n-1}(\xi) + \tfrac{1}{2} H_{n+1}(\xi).$$

and introducing this into [*149*] we obtain, as a consequence of the orthogonality of the wave functions, $\bar{x} = 0$. This result stems from the symmetry of the square of the wave function relative to the origin of coordinates.

The mean value of the square of the displacement is

$$\overline{x^2} = \frac{1}{\beta}\frac{1}{2^n n! \sqrt{\pi}} \int\limits_{-\infty}^{+\infty} H_n(\xi)\xi^2 H_n(\xi) e^{-\xi^2} d\xi. \qquad \ldots [1.150]$$

By using the recurrence formula:

$$\begin{aligned} \xi^2 H_n &= n\xi H_{n-1}(\xi) + \tfrac{1}{2}\xi H_{n+1}(\xi) \\ &= n\left[(n-1)H_{n-2} + \tfrac{1}{2}H_n\right] + \tfrac{1}{2}\left[(n+1)H_n + \tfrac{1}{2}H_{n+2}\right], \end{aligned}$$

we find that in [*150*] there remains only the integral:

$$\overline{x^2} = \frac{(n+\frac{1}{2})}{\beta 2^n n! \sqrt{\pi}} \int\limits_{-\infty}^{+\infty} [H_n(\xi)]^2 e^{-\xi^2} d\xi = \frac{n+\frac{1}{2}}{\beta}. \qquad \ldots [1.151]$$

Hence, using [*134*] and [*145*], we obtain

$$E = f\overline{x^2}. \qquad \ldots [1.152]$$

This equation corresponds to the formula [*132*] of classical mechanics, according to which the mean square displacement $\overline{x^2} = \frac{1}{2}A^2$.

7. Particle in a Spherically Symmetrical Field

Schrödinger's Equation in Polar Coordinates

In a number of problems we consider systems in which, as with the isotropic harmonic oscillator, the potential energy depends only on the radius vector r: $U = U(r)$.

The potential field in such systems has spherical symmetry. One of these problems is that of the scattering of particles by spherically symmetrical barriers; another is that of particles in spherically symmetrical potential wells. A problem of fundamental significance in atomic theory is that of the motion of an electron in the spherically symmetrical Coulomb field of the nucleus. Here the potential energy varies hyperbolically: $U(r) = -\frac{Ze^2}{r}$, where Z is the charge on the nucleus. The solution of this problem (Ch. 2, §3) is used, in particular, in the theory of the Periodic Classification of the elements and in the theory of directed bonds. The spherical symmetry of the field enables us to separate the wave function into radial and angular factors. The angular part of the wave function is uniquely determined by the symmetry of the field and does not depend on the form of the potential energy $U(r)$. The solution of problems involving spherically symmetrical systems is simplified by transforming from Cartesian to polar coordinates (r, θ, ϕ) with (Fig. 14) origin at the centre of the system.

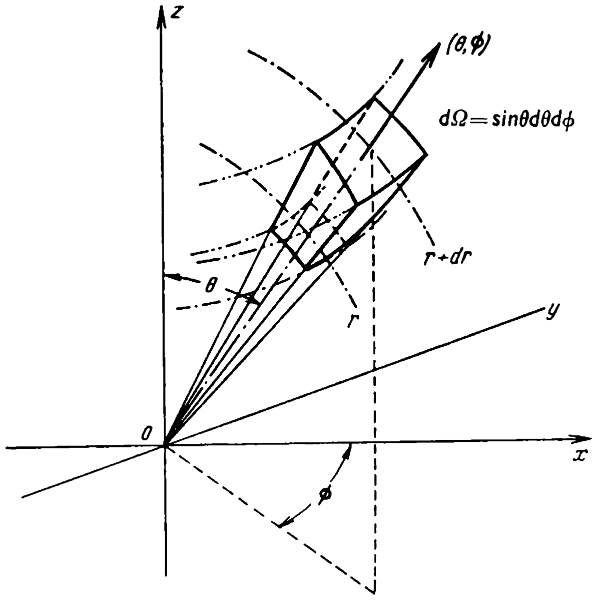


FIG. 14. Polar coordinates.

Schrödinger's equation in polar coordinates takes the form

$$\nabla^2_{r,\theta,\phi}\,\psi(r,\theta,\phi)+k^2(r)\psi(r,\theta,\phi) = 0, \qquad \ldots [1.153]$$

where $k(r)$ is given by an expression analogous to [*46*], but depending on r:

$$k(r) = \frac{2\pi}{h}\sqrt{2\mu[E-U(r)]}. \qquad \ldots [1.154]$$

Here μ is the mass of the particle. Laplace's operator in polar coordinates, as in Cartesians ($\nabla_{x,y,z} = \nabla_x+\nabla_y+\nabla_z$), is the sum of three Laplacian operators—one radial and two angular:

$$\nabla^2_{r,\theta,\phi} = \nabla^2_r+\nabla^2_\theta+\nabla^2_\phi$$

$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \qquad \ldots [1.155]$$

Note the difference in the regions of variation of polar coordinates:

$$0 \leqq r \leqq \infty, \quad 0 \leqq \theta \leqq \pi, \quad 0 \leqq \phi \leqq 2\pi.$$

The spherical coordinates θ, ϕ vary periodically.

The solutions of the wave equation [*153*] must be periodic in relation to the coordinates θ, ϕ with periods π and 2π respectively. The stipulation that a function be periodic in relation to any coordinate is equivalent to imposing boundary conditions, so the form of the wave function which depends on the coordinates θ, ϕ is uniquely determined by the spherical symmetry of the field.

Separation of Variables

Corresponding to the general method of solution for differential equations of the type under consideration, our method will be to represent the wave function as the product of three functions:

$$\psi(r, \theta, \phi) = NR(r)\,\Theta(\theta)\,\Phi(\phi), \qquad \ldots [1.156]$$

each of which depends only on one variable; N is a normalising factor. Substituting [*156*] into [*153*], dividing the whole equation by $R\Theta\Phi$ and transferring the term depending on ϕ to the R.H.S., we obtain, after multiplying by $r^2 \sin^2\theta$,

$$r^2 \sin^2\theta \left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial R}{\partial r}\right)\frac{1}{R}+k^2(r)\right]+\sin\theta \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)\frac{1}{\Theta} = -\frac{\partial^2\Phi}{\partial\phi^2}\cdot\frac{1}{\Phi}. \qquad \ldots [1.157]$$

The L.H.S. of [*157*] depends only on r and θ, and the R.H.S. only on ϕ. This is possible only if each side of the equation is equal to a constant, which we shall call m^2. The R.H.S. of [*157*] leads to the equation for Φ:

$$\frac{d^2\Phi}{d\phi^2} = -m^2\Phi. \qquad \ldots [1.158]$$

The variables can also be separated on the L.H.S. of [*157*] and we can represent it in the form:

$$r^2\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)\frac{1}{R}+k^2(r)\right] = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)\frac{1}{\Theta}+\frac{m^2}{\sin^2\theta}.$$

This equation is possible provided each of its sides has a constant value, which we shall call λ. Hence we obtain the equations for Θ and R:

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\left(\lambda-\frac{m^2}{\sin^2\theta}\right)\Theta = 0, \qquad \ldots [1.159]$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{\partial r}\right)+\left[k^2(r)-\frac{\lambda}{r^2}\right]R = 0. \qquad \ldots [1.160]$$

So equation [*153*] has broken down into three equations [*158-160*], each of which determines one of the functions Φ, Θ and R. These equations involve two constant parameters m and λ. The function $k(r)$, which determines the type of potential energy and expresses the special character of the problem, enters only into the third equation.

The normalisation condition in Cartesian coordinates has the form:

$$\int_0^\infty |\psi(x,y,z)|^2\,d\tau = \int_0^\infty\int_0^\infty\int_0^\infty |\psi(x,y,z)|^2\,dx\,dy\,dz = C,$$

where $d\tau$ is the element of volume. In polar coordinates $d\tau = r^2 \sin\theta dr d\theta d\phi$, so the normalisation condition can be written in the form:

$$\int_0^\infty |R(r)|^2 r^2 dr \int_0^\pi |\Theta(\theta)|^2 \sin\theta\, d\theta \int_0^{2\pi} |\Phi(\phi)|^2 d\phi = C.$$

This can be represented as three normalising conditions:

$$N_r^2 \int_0^\infty |R|^2 r^2 dr = N_\theta^2 \int_0^\pi |\Theta|^2 \sin\theta\, d\theta = N_\phi^2 \int_0^{2\pi} |\Phi|^2 d\phi = C^{\frac{1}{3}}. \quad \ldots [1.161]$$

The Spherical Part of the Wave Function. Legendre's Equation. Spherical Functions

The solutions of equation [*158*] are the exponential functions:

$$\Phi_m(\phi) = N_\phi e^{im\phi}.$$

The requirement that Φ_m be periodic with period 2π determines the possible values of the parameter m: these are

$$m = 0, \pm 1, \pm 2, \pm \ldots \qquad \ldots [1.162]$$

The normalisation condition [*161*] for $C = 1$ gives $N_\phi = \frac{1}{\sqrt{2\pi}}$. Consequently the normalised functions $\Phi_m(\phi)$ will be

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}, \quad m = 0, \pm 1, \pm 2, \pm, \ldots \qquad \ldots [1.163]$$

We shall transform the equation for $\Theta(\theta)$ by introducing the new variable:

$$x = \cos\theta \qquad \ldots [1.164]$$

with range of variation $-1 \leqq x \leqq 1$.
We write

$$y(x) = \Theta(\theta). \qquad \ldots [1.165]$$

The function $y(x)$ satisfies Legendre's associated equation:

$$(1-x^2)y'' - 2xy' + \left(\lambda - \frac{m^2}{1-x^2}\right)y = 0, \qquad \ldots [1.166]$$

which in the particular case $m = 0$ becomes Legendre's equation:

$$(1-x^2)y'' - 2xy' + \lambda y = 0. \qquad \ldots [1.167]$$

Legendre's equation is soluble for strictly defined values of the parameter λ, namely:

$$\lambda = l(l+1), \text{ where } l = 0, 1, 2, \ldots \qquad \ldots [1.168]$$

The solutions of equation [*167*] are the Legendre polynomials $P_l(x)$, where l is the degree of the polynomial:

$$P_l(x) = \frac{1}{2^l \cdot l!} \frac{d^l}{dx^l} (x^2-1)^l. \qquad \ldots [1.169]$$

A solution of equation [*166*] is possible only for values of the parameter λ which satisfy the condition:

$$\lambda = (k+|m|)(k+|m|+1), \text{ where } k = 0, 1, 2 \ldots, \qquad \ldots [1.170]$$

analogous to [*168*]. In [*170*] $|m|$ is the modulus of the number m, which, like k, is an integer. Writing

$$k+|m| = l,$$

we reduce [*170*] to the form [*168*]. Since k and l are positive numbers, the modulus of the number m cannot exceed l, $|m| \leqq l$. The conditions for solubility of the equations under consideration determine the values of the constant parameters λ and m:

$$\left.\begin{array}{l} \lambda = l(l+1); \quad l = 0, 1, 2, \ldots \\ m = 0, \pm 1, \pm 2, \ldots, \pm l. \end{array}\right\} \qquad \ldots [1.171]$$

The solutions of equation [*166*] are the associated Legendre polynomials:

$$y(x) = P_l^{|m|}(x) = (1-x^2)^{\frac{|m|}{2}} \frac{d^{|m|}}{dx^{|m|}} P_l(x), \qquad \ldots [1.172]$$

where l is the degree and $|m|$ the order of the polynomial. The following relationships hold for Legendre polynomials:

$$\left.\begin{array}{l} \displaystyle\int_{-1}^{+1} P_l^m(x) P_k^m(x)\,dx = 0 \qquad (k \neq l), \\ \displaystyle\int_{-1}^{+1} |P_l^m(x)|^2\,dx = \left(\frac{2}{2l+1}\right)\frac{(l+m)!}{(l-m)!} \qquad (k = l). \end{array}\right\} \qquad \ldots [1.173]$$

The first relationship shows that the polynomials are orthogonal, and the second gives the condition for normalisation. Using the latter condition and returning to the functions $\Theta(\theta)$, we obtain the normalised values of these functions:

$$\Theta_{lm}(\theta) = \sqrt{\left(\frac{2l+1}{2}\right)\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta). \qquad \ldots [1.174]$$

Combining the functions $\Phi(\phi)$ and $\Theta(\theta)$, we obtain the expression for the spherical part of the wave function:

$$Y(\theta, \phi) = N_{\theta,\phi}\Theta(\theta)\Phi(\phi). \qquad \ldots [1.175]$$

The functions $Y(\theta, \phi)$ are called *Spherical Harmonics*. When expanded, the normalised spherical harmonics may be written in the form:

$$Y(\theta,\phi) = Y_{lm}(\theta,\phi) = \sqrt{\left(\frac{2l+1}{4\pi}\right)\frac{(l-|m|)!}{(l+|m|)!}}\, P_l^{|m|}(\cos\theta)\, e^{im\phi}, \qquad \ldots [1.176]$$

where $P_l^{|m|}(\cos\theta)$ is the associated Legendre polynomial defined by formula [*172*], and l and m are integers defined by the relationships [*171*]. Putting [*175*] into [*156*], we have the wave function for a central system in the form:

$$\psi(r,\theta,\phi) = N_r Y_{lm}(\theta,\phi) R(r), \qquad \ldots [1.177]$$

where the spherical part of the wave function, $Y_{lm}(\theta, \phi)$, is uniquely determined by the spherical functions.

Quantisation of the Angular Momentum and its Projection

In classical mechanics, for closed systems the angular momentum M, and its projection M_z, as well as the total energy E, are integrals of the motion and conserve a constant value. In quantum mechanics the angular momentum and its projection are quantised like the energy of the system. To the angular momentum and its projection there correspond operators $\boldsymbol{M}$ and $\boldsymbol{M}_z$, the expressions for which are obtained from the corresponding formulae of classical mechanics by replacing the momenta p_k by operators $\boldsymbol{p}_k$ in accordance with formula [*5*]. The application of this rule leads to the following expressions for the angular momentum and its projection:

$$\boldsymbol{M}^2 = -\frac{h^2}{4\pi^2}\nabla^2_{\theta\phi}; \quad \boldsymbol{M}_z = -i\frac{h}{2\pi}\frac{\partial}{\partial\phi}. \qquad \ldots [1.178]$$

By operating with these operators on the wave function, we obtain the equations for the angular momentum and its projection:

$$\boldsymbol{M}^2\psi = M^2\psi; \quad \boldsymbol{M}_z\psi = M_z\psi, \qquad \ldots [1.179]$$

which are analogous to Schrödinger's equation $\boldsymbol{H}\psi = E\psi$. The conditions for solubility of equations [*179*] are the same as for the spherical functions [*171*]. In fact, for the projection of the angular momentum, using the expression for the operator given in [*178*] and the expression for the wave function from [*156*], we find

$$\boldsymbol{M}_z\psi(r,\theta,\phi) = -i\frac{h}{2\pi}\frac{\partial\psi}{\partial\phi} = -i\frac{h}{2\pi}R\Theta\frac{\partial\Phi}{\partial\phi}. \qquad \ldots [1.180]$$

The value of the last derivative, $im\Phi$, is obtained by differentiating [*163*]. Substituting this value into [*180*], we find

$$\boldsymbol{M}_z\psi = \frac{h}{2\pi}m\psi = M_z\psi,$$

whence we have for the projection of the angular momentum:

$$M_z = \frac{h}{2\pi}m; \quad m = 0, \pm 1, \pm 2, \ldots, \pm l. \qquad \ldots [1.181]$$

The projection of the angular momentum can take only values which are integral multiples of $\frac{h}{2\pi}$. Formula [*181*] expresses the *space quantisation* condition, which is confirmed by experiments on the scattering of beams of atoms in a nonhomogeneous magnetic field (Stern-Gerlach). Using expression [*178*] for the operator and carrying out analogous calculations, we find the allowed values of the angular momentum to be:

$$M_l^2 = \frac{h^2}{4\pi^2} l(l+1), \quad l = 0, 1, 2, \dots \qquad \dots [1.182]$$

The Equation for the Radial Part

Equation [*160*], when expanded, takes the form

$$R'' + \frac{2}{r} R' + \left[k^2(r) - \frac{\lambda}{r^2} \right] R = 0. \qquad \dots [1.183]$$

which can be simplified by substituting

$$R(r) = \frac{\rho(r)}{r}. \qquad \dots [1.184]$$

Introducing the values of $k(r)$ and λ from [*154*] and [*171*], we obtain

$$-\frac{h}{8\pi^2\mu} \rho'' + \frac{h^2 l(l+1)}{8\pi^2 \mu r^2} \rho + U(r)\rho = E\rho. \qquad \dots [1.185]$$

As regards the form of the curve $U(r)$, we shall make the following assumptions: Let $U(r)$ near the centre of the system increase no more rapidly than

$$\underset{r \to 0}{U(r)} = \frac{b}{r^\beta}, \quad \text{where} \quad \beta < 2 \ (b = \text{const}); \qquad \dots [1.186]$$

Further, at large distances from the centre let $U(r)$ tend to a constant value, which we shall take as zero:

$$\underset{r \to \infty}{U(r)} = c = 0. \qquad \dots [1.187]$$

Conditions [*186*] and [*187*] are satisfied, in particular, by the energy of the Coulomb interaction between point charges $\pm \frac{z_1 z_2 e^2}{r}$. For like charges, the potential energy has the form of a spherically symmetrical hyperbolic barrier (Fig. 15*a*) (the problem of the scattering of positively charged particles—protons, α-particles etc.—by nuclei), and for unlike charges it has the form of a spherically symmetrical potential well (Fig. 15*b*) (the problem of the hydrogen-like atom and the problem of the scattering of electrons by nuclei).

Let us consider first the solutions of equations [*185*] and [*183*] for two critical regions: (I) near the centre of the system ($r \to 0$) and (II) at a great distance from the centre ($r \to \infty$), and then join these solutions together.

I. Solution for $r \to 0$. Expand $\rho(r)$ in a power series in r and then discard all but the first term:

$$\rho(r) = Ar^\gamma \quad (A = \text{const}), \qquad \dots [1.188]$$

where γ is the index of the lowest term in the expansion. Substituting [*188*] and [*186*] into equation [*185*], we obtain

$$[\gamma(\gamma-1)-l(l+1)]\,r^{\gamma-2}+br^{\gamma-\beta}+\Delta = 0, \qquad \ldots [1.189]$$

where Δ stands for terms of higher order. Since we supposed that $\beta<2$, it follows that the lowest term in [*189*] will be the first. Discarding all terms of higher order, we obtain the equation

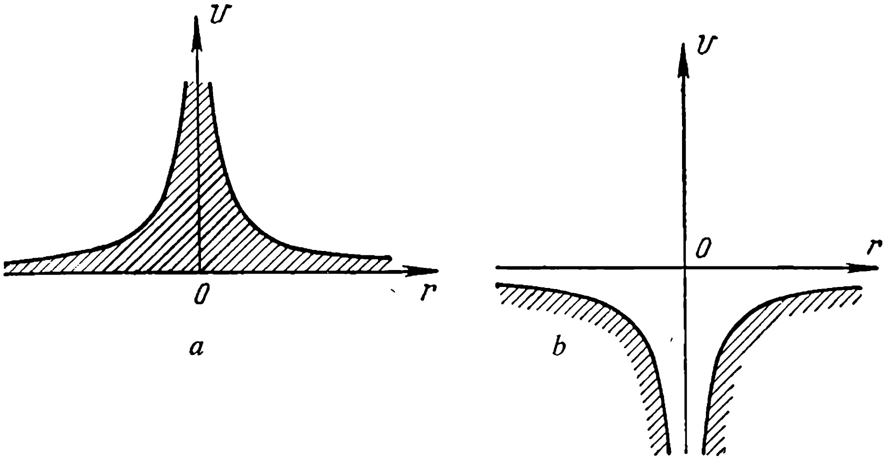


FIG. 15. Spherically symmetrical potential: (*a*) barrier, (*b*) well.

$$\gamma(\gamma-1) = l(l+1), \qquad \ldots [1.190]$$

which determines the values of γ for which the function [*188*] is a solution of equation [*185*]. Equation [*190*] has two roots:

$$\gamma_1 = l+1 \quad \text{and} \quad \gamma_2 = -l. \qquad \ldots [1.191]$$

The values found for γ determine two particular solutions of equation [*185*]. Transforming back to the radial function, we write the general solution near the centre of the system in the form

$$R_1(r) = A_1r^l+B_1r^{-(l+1)} = A_1r^l, \qquad \ldots [1.192]$$

where A_1 and B_1 are constants. The second term in [*192*] diverges for small r, so we assume

$$B_1 = 0.$$

II. Solution for $r\to\infty$. For $r\to\infty$ equation [*185*] simplifies and transforms into

$$\rho''(r)+k^2\rho = 0 \qquad \ldots [1.193]$$

with a constant value of k:

$$k = \frac{2\pi}{h}\sqrt{2\mu E}, \qquad \ldots [1.194]$$

analogous to equation [*45*]. As in the solution of the latter equation, two cases must be distinguished: (1) $E>0$ and (2) $E<0$.

(1) $E>0$. The solution of equation [*193*] is analogous to that of [*48*]. The radial function represents a superposition of two travelling spherical waves, one convergent, the other divergent:

$$R_2(r) = A_2\frac{e^{-ikr}}{r}+B_2\frac{e^{ikr}}{r}. \qquad \ldots [1.195]$$

(2) $E<0$. In this case k is imaginary. We write

$$k' = ik = \frac{2\pi i}{h}\sqrt{2\mu E}, \qquad \ldots [1.196]$$

where k' is real. The general solution of equation [*193*] is analogous to the solution for ψ_2 in [*72*] and represents the sum of two exponential functions. The solution for the radial function has the form

$$R_2(r) = A_2'\frac{e^{-k'r}}{r}+B_2'\frac{e^{k'r}}{r} = A_2'\frac{e^{-k'r}}{r}. \qquad \ldots [1.197]$$

The second term in [*197*] diverges for large r, so we assume

$$B_2' = 0.$$

The wave function is continuous in the whole range of variation of r from 0 to ∞. The solutions $R_1(r)$ and $R_2(r)$ must pass over smoothly into one another for intermediate values of r. We shall write the continuity conditions for the wave function and its derivative in the form

$$R_1 \rightleftarrows R_2, \quad R_1' \rightleftarrows R_2'. \qquad \ldots [1.198]$$

These conditions, together with the normalisation requirement, give three equations for the coefficients of the asymptotic solutions. In the case $E>0$, which corresponds to the low barrier or the weak binding, the equations above enable us to determine the three unknown amplitudes A_1, A_2, B_2. The solutions of these equations are not subject to any additional limitations, so the spectrum of the particle is continuous.

In the case $E<0$, corresponding to the strong binding, the position is fundamentally different, since there is an additional condition $B_2' = 0$. The continuity relationships [*198*] change into a system of two homogeneous equations of the form

$$a_1A_1+b_1A_2' = 0; \quad a_2A_1+b_2A_2' = 0. \qquad \ldots [1.199]$$

Equations [*199*] are compatible provided the determinant formed from the coefficients in the equations vanishes:

$$\begin{vmatrix} a_1, & b_1 \\ a_2, & b_2 \end{vmatrix} = 0. \qquad \ldots [1.200]$$

Since the coefficients b_1 and b_2 depend on k', that is on the particle energy, [*200*] represents the equation for the energy E and contains the quantisation conditions. The particle energy in a spherically symmetrical well is quantised in the same way as in the case of a parabolic well or a potential box.

8. Particle in a Periodic Field

Particle in a Periodic Field. Weak and Strong Binding

In the majority of cases the stable state of a solid body is the crystalline state, which is characterised by a regular, periodic arrangement of the atoms. Any periodic field is thus a sequence of identical potential wells and barriers (Fig. 16). The question of the motion of particles in a periodic field constitutes one of the important problems of quantum mechanics as applied to the solid state. The fundamental peculiarities of this problem may be understood in the light of the results obtained from a consideration of the passage of a particle through separate potential wells and barriers. For a periodic field, as for a single well, two extreme cases must be distinguished—weak and strong binding.

With strong binding the particle energy is considerably less than the height of the barriers, so the particles are trapped by the wells and their wave functions represent standing waves, confined by the dimensions of the individual wells. The particle energy is quantised and the energy spectrum is discrete. For low particle energies the barriers which separate the wells are practically impenetrable. Particles in neighbouring wells do not interact with one another; their energy levels are identical and the whole system of particles must be regarded as degenerate.

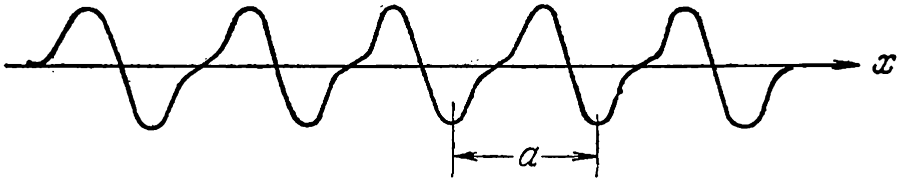


FIG. 16. Periodic potential field.

With weak binding, the particle energy exceeds the height of the barriers so the particles pass freely over the barriers, experiencing periodic velocity variations. The wave functions of the particles represent travelling waves whose amplitudes are periodic functions with the same period as the field. The energy spectrum of the particles is, in the limit, continuous. Strengthening the binding causes strengthening of the periodic perturbations in the motion of each of the particles and makes breaks appear in the energy spectrum, separating it into distinct bands. Further strengthening of the binding leads to the particles being captured and the energy bands degenerating into line levels. The transitional case can be considered from the point of view of the penetrability of the barriers. With strong binding, as was observed above, the barriers are impenetrable. If the particle energy is increased till it approaches the peak energy of the barrier, the barrier becomes more penetrable. The particles begin to pass through into neighbouring wells, and this results in the removal of the degeneracy and splitting of the energy levels.

Solving the problem of a particle moving in a periodic field enables us to obtain the complete energy spectrum of the particle, with the transition from strong to weak binding. This spectrum makes it possible to give a single description of the behaviour of the different electrons belonging to atoms in crystals. The case of strong binding occurs in bound electrons in atomic structures, and the case of weak binding is exemplified by valence electrons, including the free electrons in metals. The picture we have set forth is also adequate to describe the oscillatory motion of atomic nuclei in crystals, a case of strong binding.

Particle in a Sinusoidal Field. Mathieu's Equation

Consider a particle moving along the x-axis in an arbitrary periodic field, the length of the period being a (Fig. 16). We expand the potential energy of the particle in a Fourier series:

$$U(x) = \sum_{n=-\infty}^{+\infty} U_n \cos 2\pi n \frac{x}{a}. \qquad \ldots [1.201]$$

Take one of the terms in the series and consider the motion of a particle in a sinusoidal field. As origin we take one of the sinusoidal minima. The potential energy of such a field may be written in the form

$$U(x) = U_0 - U_0 \cos 2\pi \frac{x}{a}. \qquad \ldots [1.202]$$

Introducing this value for $U(x)$ into Schrödinger's equation [*19*], we obtain

$$\psi''(x) + \frac{8\pi^2 m}{h^2}\left(E - U_0 + U_0 \cos 2\pi \frac{x}{a}\right)\psi(x) = 0. \qquad \ldots [1.203]$$

Introduce the new variable

$$\xi = 2\pi \frac{x}{a} \qquad \ldots [1.204]$$

and write

$$\eta = \frac{8a^2 m(E-U_0)}{h^2}, \quad \gamma = \frac{8a^2 m U_0}{h^2}. \qquad \ldots [1.205]$$

Equation [*203*] can then be written in the form

$$\psi''(\xi) + (\eta + \gamma \cos \xi)\psi(\xi) = 0. \qquad \ldots [1.206]$$

Equation [*206*] is Mathieu's equation. As was shown by Flocquet, the general solution of this equation has the form

$$\psi(\xi) = A\Phi(\xi)e^{ik\xi} + B\Phi(-\xi)e^{-ik\xi}. \qquad \ldots [1.207]$$

Here A and B are constants, $\Phi(\xi)$ is a periodic function of ξ with period 2π, and k is a quantity depending on the parameters η and γ of the equation. The solution of [*207*] is analogous to the solution of [*48*] for a free particle, the difference being connected with the function $\Phi(\xi)$, which may be regarded as a variable amplitude, varying in conformity with the periodicity of the field. The solution of [*207*] represents a superposition of two travelling waves or the sum of two exponential functions, according as k is real or imaginary. The exponential solutions lead to infinitely large values of the wave function, and so we shall henceforth consider only periodic solutions, that is, solutions with a real value of k. Periodic solutions do not exist for arbitrary values of η and γ: a given value of η is possible only for a definite value of γ. To each value of k there corresponds a definite curve in the (η, γ) plane (Fig. 17). For small $\gamma \to 0$ equation [*206*] transforms into the equation for a free particle. Its solutions are sinusoidal waves of the type [*47*]. The wave number is given by the relationship

$$k = \sqrt{\eta}, \qquad \ldots [1.208]$$

which means that it is a parabolic function of the energy. For $\gamma = 0$ the energy E and the parameter η vary continuously. There is a corresponding continuous variation of k from 0 to ∞. The lines $\gamma = \gamma(\eta)$, corresponding to $k = \text{const.}$, will originate on the x-axis and, as γ increases, they merge with the curves corresponding to integral values, $k = 0, 1, 2, 3, \ldots, n, \ldots$ The shaded regions in Fig. 17 correspond to real k, and the unshaded to imaginary k. The region of the diagram to the left of the line $\eta = -\gamma$ corresponds to energies E less than the minima of the potential wells, and is of no interest to us. The region included between the lines $\eta = -\gamma$ and $\eta = +\gamma$ corresponds to values of the particle energy in the range $0 \leqq E \leqq 2U_0$, lying within the limits of the potential well, and represents strong binding. To the right of the line $\eta = +\gamma$ lies the region of weak binding.

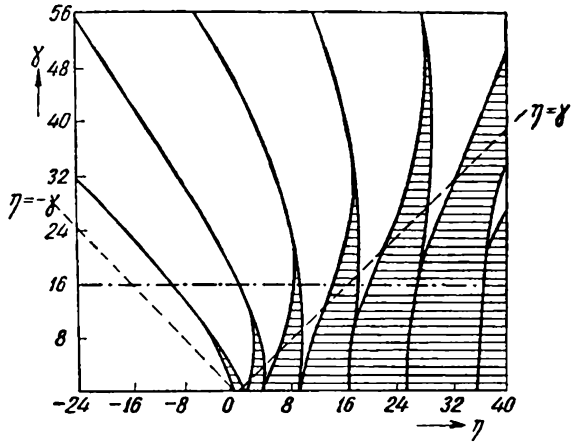


FIG. 17. Regions of periodic and non-periodic solutions of equation *1.206*.

Near the bottom of the well the potential energy may be expanded as a series, and for small ξ, approximated by the parabola $U(\xi) = U_0 - U_0 \cos \xi \approx \frac{1}{2} U_0 \xi^2$. For a sufficiently deep well the sequence of energy levels corresponds to the levels of the quantised oscillator $E_n = h\nu \, (n + \frac{1}{2})$. Near the tops of the barriers the levels spread out more, but nevertheless still remain separated by forbidden regions. In a periodic field, motion is not possible for all values of the particle energy, even in those cases when the particle energy exceeds the barrier height. The possible values of the particle energy are shown in Fig. 18 by shaded bands, narrowing to lines deeper down in the wells.

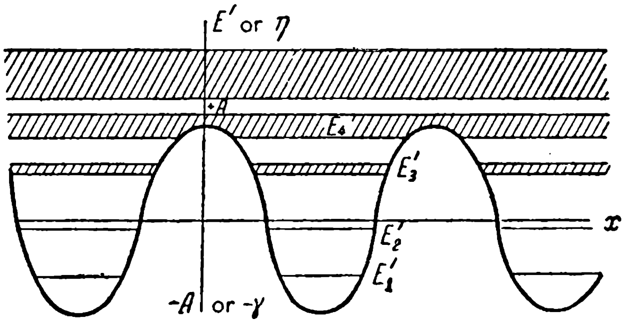


FIG. 18. Possible values of the particle energy in a periodic field.

We now find the relationship between the wave number k and the parameter η (the particle energy). For small $\gamma \ll 1$, as is shown by [*208*], this relationship is expressed by a parabola (Fig. 19*a*) cut into separate parts by vertical ordinates corresponding to the integral values of k. For greater barrier heights, for example for $\gamma = 16$, as can be seen from Fig. 17, the energy discontinuities become greater (Fig. 19*b*). From the first relationship in [*205*] we have

$$E = U_0 + \frac{h^2}{8m} \frac{\eta}{a^2}. \qquad \ldots [1.209]$$

Using [*208*], we obtain for small γ

$$E \simeq U_0 + \frac{h^2}{2m}\left(\frac{k}{2a}\right)^2 = U_0 + \frac{h^2}{2m}\left(\frac{n}{2a}\right)^2. \qquad \ldots [\mathit{1.210}]$$

Comparing [210] with [52] and [50], gives

$$\frac{1}{\lambda} = \frac{k}{2a}. \qquad \dots [1.211]$$

The energy discontinuities correspond to integral k and the condition for discontinuities to occur can be written in the form

$$n\lambda = 2a. \qquad \dots [1.212]$$

Formula [212] represents the condition for diffraction maxima for wave motion along a one-dimensional lattice. The appearance of energy discontinuities is explained by diffraction of the particles by the periodic structure. For a three-dimensional lattice [212] generalises into the Bragg-Wulff formula. The greater the energy discontinuities, the greater is the scattering power of the lattice planes. This latter is measured by the structural factor well known in structural analysis.

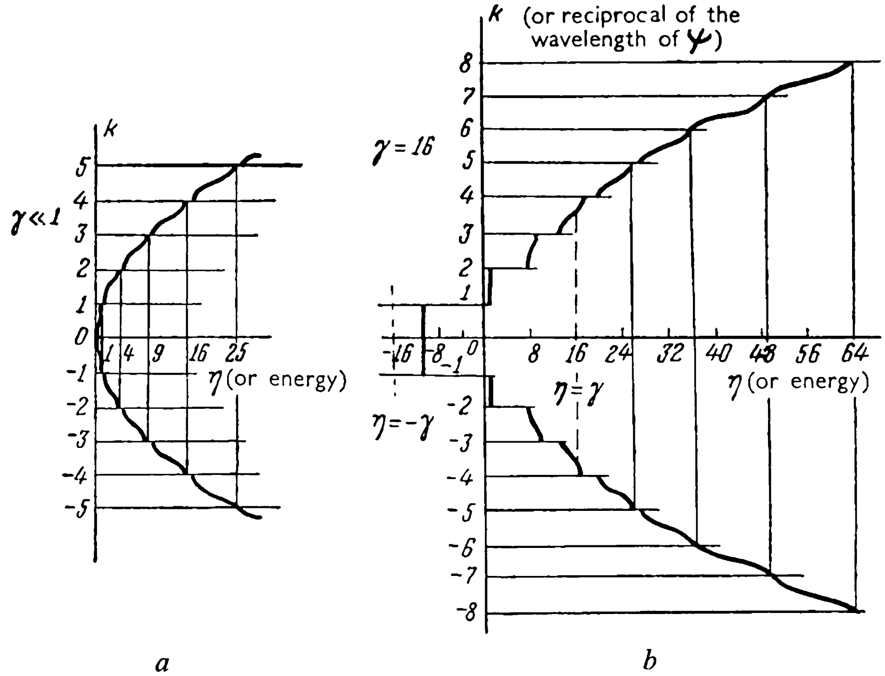


FIG. 19. Dependence of η on γ for (a) weak and (b) strong binding.

Particle in a One-dimensional Periodic Field. Hill's Equation

For a periodic field of arbitrary shape with potential $U(\xi)$ and period 2π, Mathieu's equation [206] generalises to Hill's equation:

$$\psi''(\xi)+[\eta+\gamma U(\xi)]\,\psi(\xi) = 0. \qquad \dots [1.213]$$

The general solution of equation [213] is the same as that of [206], and is given by formula [207]. The solution of Hill's equation can also be represented in graphical form. Let U_M and U_m be respectively the greatest and the least values of the potential. In the (γ, η) plane (Fig. 20) we draw the straight lines $\eta = \gamma U_m$ and $\eta = \gamma U_M$, which cut the positive half-plane into three regions. Region III corresponds to weak binding, and region II to strong binding. Region I is not realised, as it corresponds to particles with energies less than the potential minima. Fig. 21 shows the regions of real (shaded) and imaginary k for a periodic field composed of rectangular barriers and wells of width $\frac{1}{2}a$. The general character of the diagram in this case is the same as for the sinusoidal field.

Calculation of the Translational Symmetry of the Field

The general expression for the wave function of a particle in a periodic field is of the form

$$\psi(\boldsymbol{r}) = u_k(\boldsymbol{r})e^{i\boldsymbol{k}\boldsymbol{r}}, \qquad \ldots [1.214]$$

obtained by Bloch, where $\boldsymbol{k}$ is the wave vector, and $u(\boldsymbol{r})$ is a periodic function with the same periodicity as the field:

$$u(\boldsymbol{r}+\boldsymbol{a}) = u(\boldsymbol{r}). \qquad \ldots [1.215]$$

This expression results from the fundamental property of a periodic field—its *translational symmetry* (Ch. 5, §4). In formula [*215*] $\boldsymbol{a}$ is the translational displacement vector. For a three-dimensional periodic field

$$\boldsymbol{a} = n_1\boldsymbol{a}_1+n_2\boldsymbol{a}_2+n_3\boldsymbol{a}_3, \qquad \ldots [1.216]$$

where n_1, n_2, n_3 are integers, and $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ are the three non-coplanar vectors such that displacements which are parallel to, and integral multiples of them map out the three-dimensional lattice. The operation of parallel translation through a distance equal to the magnitude of the lattice vector:

$$\boldsymbol{r} \rightarrow \boldsymbol{r}+\boldsymbol{a},$$

is an invariant transformation, since it brings the system into a state which coincides with its initial state. Consequently, the wave functions $\psi(\boldsymbol{r})$ and $\psi(\boldsymbol{r}+\boldsymbol{a})$ must be the same apart from a constant factor:

$$\psi(\boldsymbol{r}+\boldsymbol{a}) = C\psi(\boldsymbol{r}), \qquad \ldots [1.217]$$

the modulus of the constant factor being unity:

$$|C| = 1. \qquad \ldots [1.218]$$

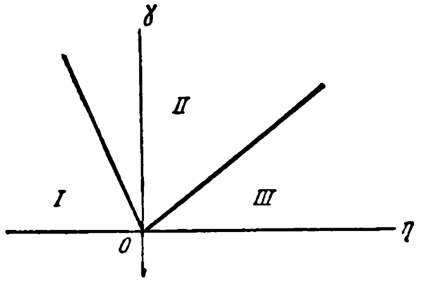


FIG. 20. The different regions in the solution of Hill's equation.

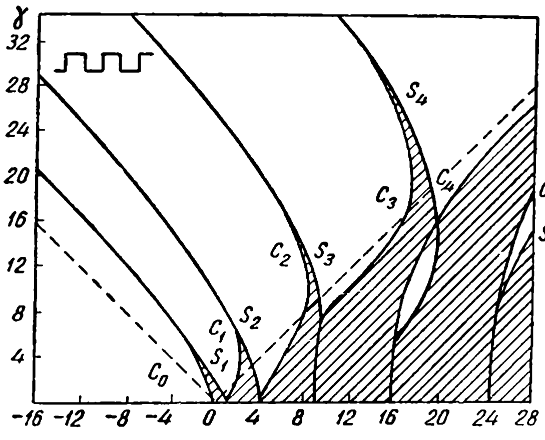


FIG. 21. Regions of periodic and non-periodic solutions of Hill's equation for square barriers.

If condition [*218*] were not satisfied, the wave function would increase or decrease without limit during multiple repetition of the operation of parallel translation.

In view of the independence of translations along different coordinate axes, it is sufficient to consider the operation of parallel translation along only one of the axes. We shall take a one-dimensional chain consisting of N lattice points. In order to exclude boundary conditions, we shall close the chain into a ring of total length $L = Na$. If we now repeat the elementary translation operation N times, which will bring us back to the initial point. Using the equation of the symmetrical transformation [*217*], we can express these transformations in the analytical form:

$$\psi(x+Na) = C^N\psi(x) = \psi(x). \qquad \ldots [1.219]$$

Whence we obtain the equation for the modulus of the wave function

$$C^N = 1, \qquad \ldots [1.220]$$

the solutions of which are the roots of unity. Equation [*220*] has N solutions:

$$C_n = e^{\frac{2\pi i n}{N}}, \quad n = 1, 2, \ldots, N. \qquad \ldots [1.221]$$

The function ψ must satisfy conditions [*219*] and [*221*] for any value of x. This is possible provided

$$\psi(x) = u_k(x)e^{\frac{2\pi i n x}{Na}}, \qquad \ldots [1.222]$$

where $u_k(x)$ is a periodic function of x with period a. We may write

$$k = \frac{2\pi n}{Na}, \qquad \ldots [1.223]$$

and then [*222*] becomes

$$\psi(x) = u_k(x)e^{ikx},$$

which is the same as [*214*].

Effect of an External Field. Effective Mass of a Particle

Let a particle situated in a periodic field be acted on by an external field. If F is the force acting on the particle, then the element of work done by the field is

$$dE = Fvdt, \qquad \ldots [1.224]$$

where v is the group velocity of the particle, given by formula [*55*]. Using the value of ω from [*21*] and formula [*55*], we find for the group velocity of the particle:

$$v = \frac{2\pi}{h}\frac{dE}{dk}. \qquad \ldots [1.225]$$

Hence

$$dE = \frac{h}{2\pi} vdk. \qquad \ldots [1.226]$$

Comparing [*224*] and [*226*], we obtain

$$\frac{dk}{dt} = \frac{2\pi}{h}F. \qquad \ldots [1.227]$$

Differentiating both sides of [*225*] with respect to t, we obtain

$$\frac{dv}{dt} = \frac{2\pi}{h}\frac{d^2E}{dk^2}\frac{dk}{dt} = \frac{d^2E}{dk^2}\cdot\frac{4\pi^2}{h^2}F. \qquad \ldots [1.228]$$

This formula establishes the connection between the acceleration and the effective force and expresses Newton's second law. For a free particle this law has the form

$$\frac{dv}{dt} = \frac{F}{m}. \qquad \ldots [1.229]$$

Comparison of [*228*] and [*229*] shows that a particle in a periodic field is accelerated by an external force as if it possessed a mass

$$m^* = \frac{h^2/4\pi^2}{d^2E/dk^2}. \qquad \ldots [1.230]$$

The quantity m^* is called the effective mass. For a free particle (sec [*52*]), the dependence of E on k is expressed by a parabola $E = \frac{h^2}{8\pi^2 m}k^2$, and it is not difficult to see that in this case $m^* = m$.

Fig. 22 shows the curve $E = E(k)$ for a particle in a periodic field. The effective mass is given by the divergence of the curvature of this curve from the curvature of the parabola: in the middle of the allowed zones the curvatures of the two curves coincide, but near the bottom and top of the zone, where the energy discontinuities appear because of the selective reflection of particles, the radius of curvature alters. The effective mass of a particle in these cases is considerably altered. The sign of the radius of curvature for states near the bottom of the zone is the same as for a free particle (positive). For a state near the upper region of the zone the sign of the curvature is negative. This means that the effective mass in this case is negative. Charged particles with negative effective mass move in electromagnetic fields like particles carrying charges of the opposite sign.

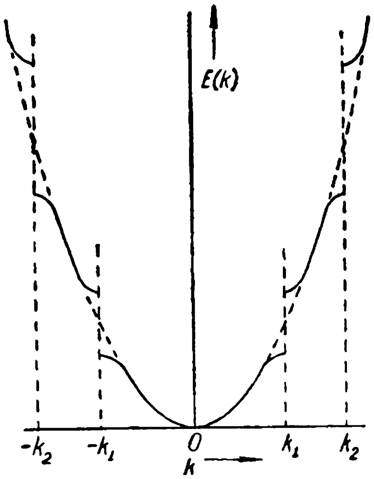


FIG. 22. The dependence of energy on momentum for a particle in a periodic field.

The electrons in crystals which occupy the upper energy levels of the zones move like positively charged particles. This quantum-mechanical effect explains the positive value of the Hall constant in certain metals and semiconductors. The absolute value of the ratio m^*/m for electrons in metals may be either greater or less than unity. For example, in palladium $m^*/m = 43$.

Problems

1. Find the connection between the phase and group velocities of de Broglie waves. *Hint*: use the relativistic expressions for the total energy and momentum of the particle.
2. Find the group velocity of de Broglie waves for a wave group of arbitrary shape.
3. Find the energy and angular momentum of a rigid rotator (diatomic molecule).

Chapter 2

ATOMIC STRUCTURE AND THE PERIODIC TABLE OF THE ELEMENTS

1. The Atomic Nucleus

The Nuclear Model of the Atom. Elementary Particles

Experiments on the scattering of α-particles, electrons and X-rays by atoms show that atoms consist of a positively charged nucleus, surrounded by a thin atmosphere of electrons. The nucleus consists of particles which are heavy compared to the electrons. These are called nucleons, and comprise protons and neutrons. Some data relating to the elementary particles which form atoms are given in Table 1.

Table 1

Data relating to Elementary Particles

Particle	Mass, in g.	Charge e, in e.s.u.	Moment	
			Mechanical, $\frac{h}{2\pi}$	Magnetic
Electron	$9{\cdot}108\times10^{-28}$	$4{\cdot}803\times10^{-10}$	$\frac{1}{2}$	$-1{\cdot}00116M_B$
Proton	$1{\cdot}672\times10^{-24}$	$4{\cdot}803\times10^{-10}$	$\frac{1}{2}$	$2{\cdot}785M_{\text{nuc}}$
Neutron	$1{\cdot}675\times10^{-24}$	0	$\frac{1}{2}$	$-1{\cdot}935M_{\text{nuc}}$

The intrinsic mechanical moments of the electron and the nucleons, P_s, are called spin moments, and are the same. The spin quantum number for these particles is $s = \frac{1}{2}$. The spin of an electron is well described by relativistic quantum theory, from which it follows that the intrinsic moment of an electron is equal to the Bohr magneton:

$$M_B = \frac{he}{4\pi\mu c} = 0{\cdot}927\times10^{-20} \text{ erg/oersted}, \qquad \ldots [2.1]$$

where e and μ are the charge and mass of the electron, $h = 6{\cdot}625\times10^{-27}$ erg-sec. is the quantum constant, and c is the velocity of light. The slight divergence of the magnetic moment of an electron from unity, $M_s = 1{\cdot}001\ M_B$, is connected with the recently discovered effect of the interaction between an electron and empty space. The gyromagnetic ratio for the electron spin is

$$\frac{M_s}{P_s} = 1{\cdot}001\,\frac{e}{\mu c}. \qquad \ldots [2.2]$$

The magnetic moments of the nucleons are considerably less than that of an electron, and are expressed in terms of nuclear magnetons:

$$M_{\text{nuc}} = \frac{he}{4\pi m_p c} = 0{\cdot}505 \times 10^{-23} \text{ erg/oersted}, \qquad \ldots [2.3]$$

where m_p is the mass of the proton. Formula [*3*] represents an extrapolation to the case of heavy particles of formula [*2*], which was derived for the electron. Consequently formula [*3*] does not describe so accurately the properties of nucleons. Measurements of the magnetic moments of the proton and neutron give coefficients differing significantly from unity.

The Structure of the Atomic Nucleus

The structure and properties of atomic nuclei, except for certain cases of nuclear isomers, are uniquely defined by the composition of the nucleus, that is, by the number Z of protons and the number N of neutrons. The number of protons determines the nuclear charge and atomic number of the corresponding atom. The mass of the nucleus, multiplied by a packing factor which differs only slightly from unity, is equal to the sum of the masses of the protons and neutrons. Since the masses of the proton and the neutron are very nearly the same, the mass of the nucleus is approximately given by the mass number A:

$$A = Z + N. \qquad \ldots [2.4]$$

The nucleons in the nuclei are united by strong nuclear forces, operating at very small distances $\sim 10^{-13}$ cm. The mean binding energy per nucleon in the nucleus is 10^7 eV, which is several million times greater than the energy of the cohesive forces between atoms (the energy of the chemical bond). The linear dimensions of nuclei are small, of the order of 10^{-12}–10^{-13} cm.; they increase in proportion to the number of nucleons. The density of matter in the nucleus is very high and is approximately constant for the different nuclei. The nuclei can be looked on as very minute drops of nuclear fluid.†

Stable nuclei are formed only for certain definite ratios of number of protons to number of neutrons. When the nuclear charge is increased, the nuclear stability at first increases slightly, but later, for $Z > 40$, diminishes because of the electrostatic repulsion of the protons. Fig. 23 shows a proton-neutron diagram on which the stable nuclei for light and heavy elements are marked. The diagram shows that the stable nuclei form a narrow locus, which for light nuclei coincides with the bisector of the angles between the coordinate axes ($N = Z$), but for medium and heavy nuclei diverges slightly from it in favour of an excess of neutrons ($N > Z$). From this follows the monotonic dependence between nuclear mass and charge:

$$A/Z \approx 2 \text{ to } 2{\cdot}5. \qquad \ldots [2.5]$$

This relationship was, in its essential principles, used by Mendeleev and enabled him to discover the periodicity of the chemical and physical properties of the elements.

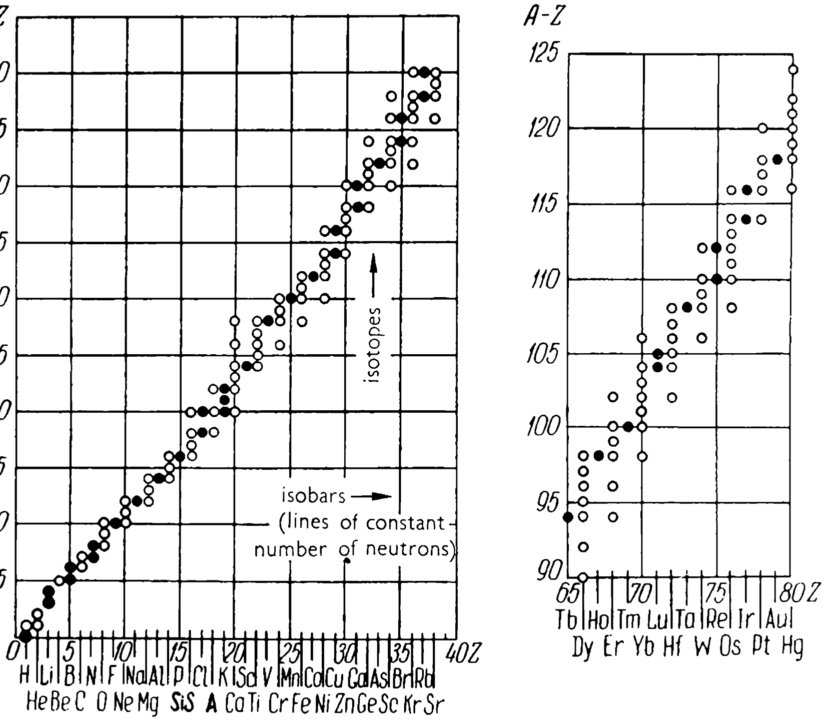
The total number of stable nuclei is about 300, and of radioactive nuclei about 500. Thus in the chemistry of atomic nuclei about 800 different composite nuclei are known and this apparently exhausts the number of possible nuclei. Nuclei with atomic numbers $Z > 100$ are extremely unstable because of the large nuclear charge and the electrostatic repulsion of the protons, and disintegrate either by fission or by a series of successive radioactive transformations. The theory of the atomic nucleus enables us to understand the reason why there is a limit to the number of elements in the periodic table.

---

† For the 'liquid drop' model of the atomic nucleus, see, for example, J. C. Slater, *Modern Physics*, McGraw-Hill, 1955.

Isotopes and the Difference between the Physical Properties of Substances with Different Isotopic Compositions

In the diagram of Fig. 23 we can distinguish between nuclear isobars (A = const.) and nuclear isotopes (Z = const.). The isobars give rise to atoms with different numbers of electrons, that is to atoms of different elements. Nuclear isotopes are atoms with the same number of electrons, possessing the same chemical, and very nearly the same physical, properties. There are 24 elements known with only a single isotope (beryllium, fluorine, sodium, aluminium and others); the remaining elements are a mixture of isotopes (tin, for example, has 11 isotopes). The atomic weights of these elements differ considerably from integral values. They depend on the percentages of the various isotopes which occur naturally in the mixtures.



a *b*

FIG. 23. Stable isotopes (*a*) of light and (*b*) of heavy elements (ratio of number of protons to number of neutrons in the atomic nuclei).

The differences between the properties of substances with different isotopic compositions are caused by the variations in the atomic masses of isotopes, since the mass affects the diffusional mobility of the free isotope and the frequency of oscillations of bound isotopes. As a result there are variations in the zero point energy $\frac{1}{2}h\nu$. The most striking isotopic effects occur in connection with the lightest atom—that of hydrogen. Three isotopes of hydrogen are known: protium, deuterium and tritium, with mass ratios 1:2:3. Table 2 gives some properties of light and heavy water.

Table 2

Some Properties of Light (H_2O) and Heavy (D_2O) Water

Water formula	Melting point, °C.	Boiling point at 1 atm., °C.	Ice density at melting point, in g./cm.^{3}	Thermal capacity at 25° C., in cal./mole	Viscosity at 25° C., in poises
H_2O	0	100	0·9176	18·00	0·891
D_2O	3·82	101·42	1·0148	20·1	1·099

Changing light hydrogen into deuterium markedly alters the boiling and melting points, density, specific heat and so on.

In the crystals of metallic hydrides, the isotope change from hydrogen to deuterium diminishes the frequency of oscillation of the atom of deuterium and diminishes the zero point energy $\frac{1}{2}h\nu$ (Ch. 3, §4, p. 101). This corresponds to the low effective temperature of a deuteride as compared with a hydride, and leads to a shortening of the interatomic distances. This effect, which is of quantum origin, is confirmed by X-ray measurements. In the cubic deuterides of uranium (UD_3), sodium (NaD) and lithium (LiD), the dimensions of the unit cells are 0·01-0·02 Å less than in the corresponding hydrides, the divergence amounting to several tenths of 1%. An analogous effect, but an order of magnitude smaller, takes place also in the case of heavier isotopes. Thus the lattice parameter of Li^7F is 0·0008 Å less than that of Li^6F: the divergence here amounts to 0·02%.

In molecular crystals, in contrast, the change from hydrogen to deuterium may lead to an increase in the intermolecular distances. Thus the unit cell of heavy ice is somewhat larger than that of light ice in the direction of the *a*-axis. The explanation is that in the water molecule itself the interatomic distance O–D is less than the distance O–H. Consequently the dipole moment of the molecule of heavy water is less, and therefore so also is the intermolecular interaction.

A considerable change is observed in the electrical properties of some ferroelectric crystals when isotopes are substituted. Thus for KD_2PO_4 the ferroelectric Curie point drops 100° C as compared with KH_2PO_4. A shift $\Delta T_c \sim 0{\cdot}01°\,K$ in the superconducting transition temperature has been observed in several isotopes of superconducting metals (tin, mercury) (Fig. 170). In this case the relationship holds: $T_c m^{1/2} = \text{const}$. The smallness of the shifts is explained by the extreme closeness of the mass ratios of the isotopes to unity, and the fact that the effects themselves are observed at very low temperatures. The isotopic effect has recently been established in the electric resistance of metals: at low temperatures the electric resistance of the heavy isotope Li^7 turned out to be greater than that of Li^6. This is in qualitative agreement with the theory of metallic conduction.

The Relative Abundance of the Various Elements

An increasing number of elements, including even the so-called rare elements, are finding practical applications. Some of them (for example technetium, promethium, plutonium) are found in nature in such minute quantities that they have to be produced artificially, but the majority of elements are obtained from natural deposits.

The question of the abundance of elements has great practical importance. Their relative abundances are different in the universe and on the earth. In the theory of how the elements originated, the first consideration is their relative cosmic abundance. The data given by analysis of different celestial objects (stellar atmospheres, interstellar matter, meteorites) are in satisfactory agreement. At the same time they differ significantly from the data obtained by analysis of the earth's crust though one must bear in mind that the earth's crust may possibly not reflect the material composition of the earth as a whole.

Table 3 gives data relating to the cosmic abundance of certain elements.

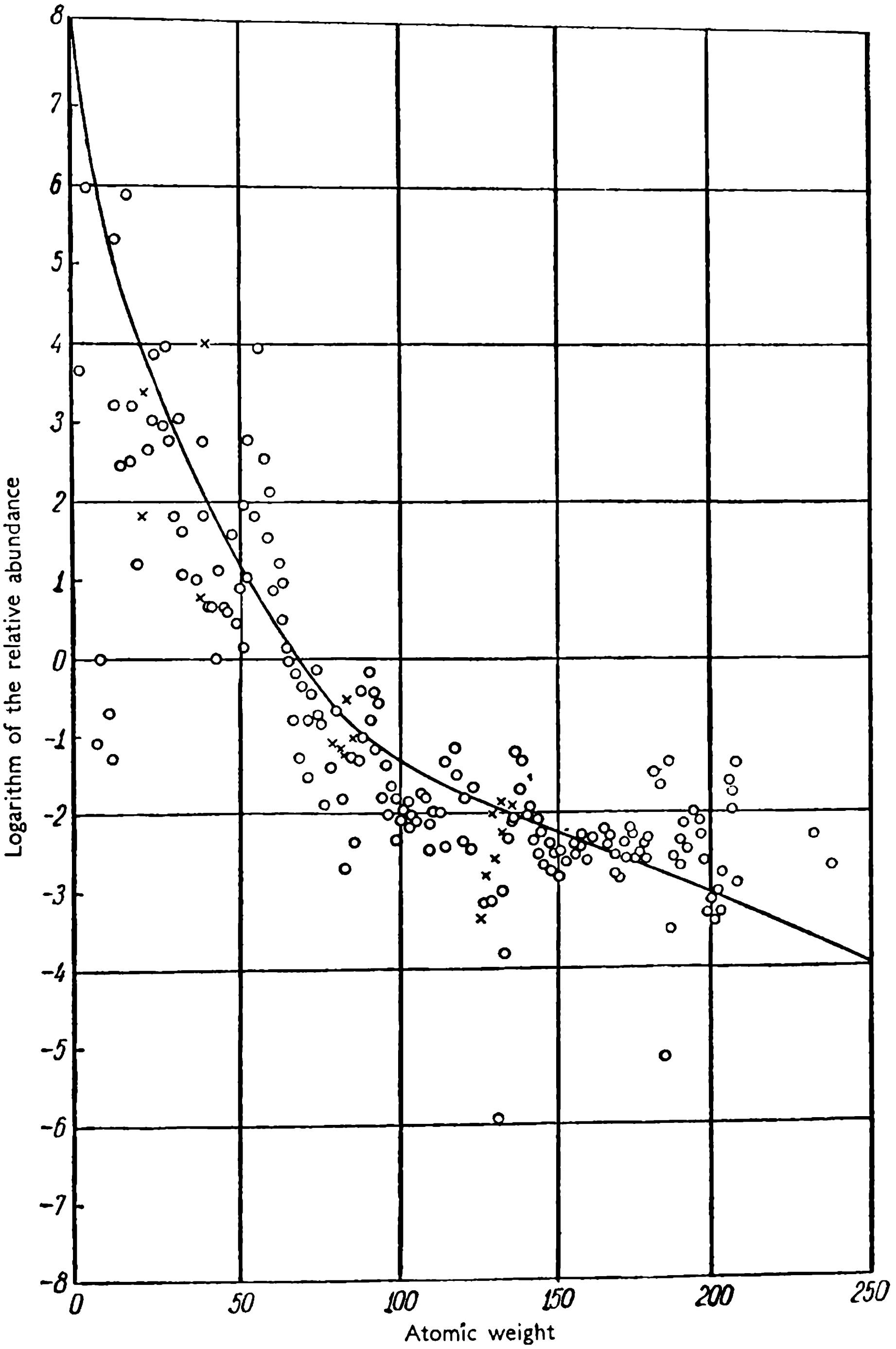


FIG. 24. Dependence of relative abundance of elements on atomic number.

Table 3

Cosmic Abundance of Certain Elements (Expressed in Terms of the Relative Number of Atoms)

Element	Atomic number Z	Atomic weight A	Abundance	Element	Atomic number Z	Atomic weight A	Abundance
H	1	1·01	$3{\cdot}5\times10^{8}$	Ni	22	58·69	2×10^{3}
He	2	4	$3{\cdot}5\times10^{7}$	Cu	29	63·57	7
Be	4	9·02	2×10^{-1}	Ga	31	69·72	5×10^{-3}
C	6	12·01	8×10^{4}	Sr	38	87·63	1×10^{-1}
O	8	16·00	1×10^{5}	Cd	48	112·41	2×10^{-2}
Si	14	28·06	1×10^{4}	Cs	55	132·91	1×10^{-2}
Cl	17	35·46	$2{\cdot}5\times10^{2}$	Pt	78	195·23	1×10^{-1}
Mn	25	54·93	1×10^{2}	Pb	82	207·21	4×10^{-3}
Fe	26	55·85	$2{\cdot}6\times10^{4}$	Th	90	231·12	1×10^{-2}
Co	27	58·94	$1{\cdot}6\times10^{2}$	U	92	238·07	3×10^{-3}

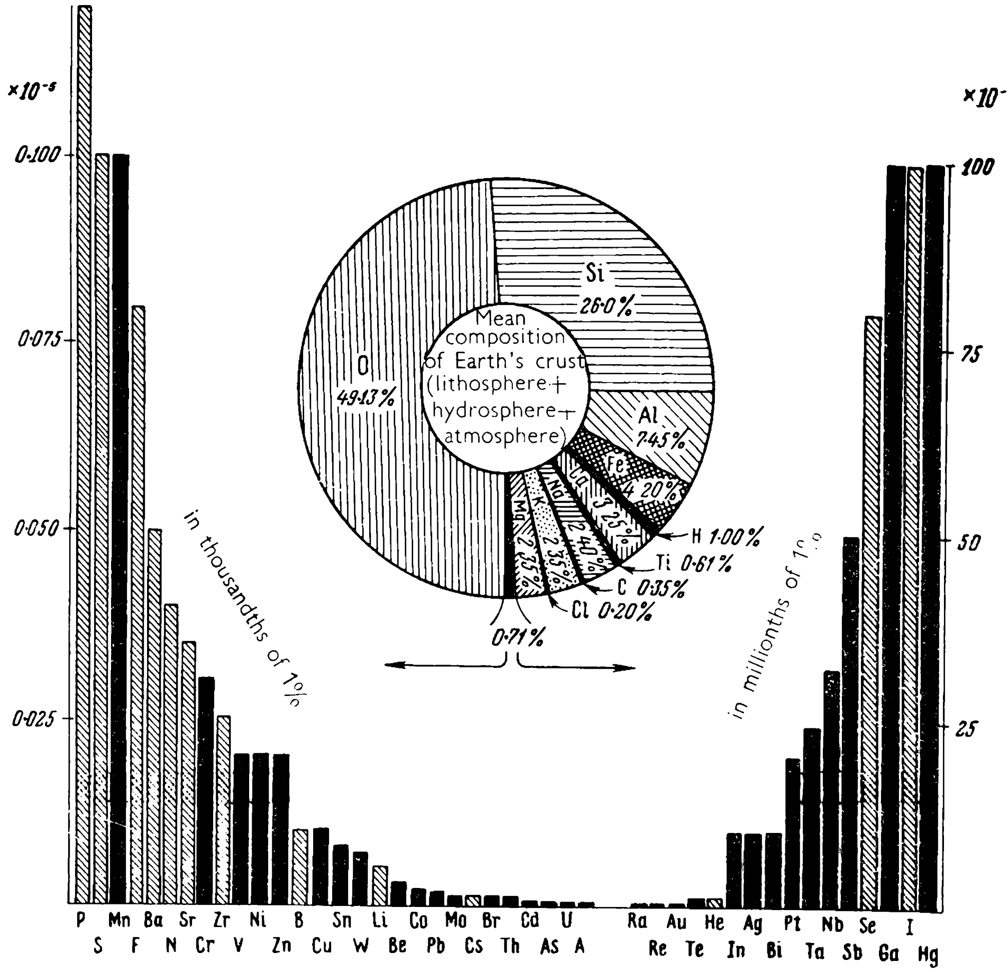


FIG. 25. The commonest elements of the Earth's crust (lithosphere, hydrosphere, and part of the atmosphere) in % weight.

The commonest element is the simplest—hydrogen. After it comes helium. The relative abundance of the elements which come after helium (lithium, beryllium, boron) drops sharply to small values. Between helium and beryllium this drop amounts to a factor $\sim 10^8$. The light elements which come next (carbon) are relatively common, about as common as, for example, silicon. With oxygen there is a sharp rise in the prevalence. Further on, the elements become rarer with increasing atomic weight. Fig. 24 gives data on the cosmic prevalence of the elements, and a mean curve. The group of relatively common elements is completed by iron.

Fig. 25 shows the prevalence of the principal elements in the upper layers of the earth's crust (conventionally to a depth of 16 km.). Only a few elements are found in large quantities in the earth's crust; these are elements with low atomic numbers, such as oxygen, silicon, aluminium etc. Twelve of these elements, as Fig. 25 shows, account for more than 99% of the mass of matter contained in the earth's crust. All the others amount to 0·71% by weight.

2. Fundamental Properties of Atoms

Discrete and Stationary Nature of the Energy Levels of Atoms

Among the important properties of atoms are the discrete and stationary nature of the energy levels. These properties appear in such experimental results as: (*a*) the line character of atomic spectra, (*b*) the discreteness of the excitation and ionisation potentials and their connection with the atomic spectra, (*c*) the elasticity of low energy collisions of atoms and ions with one another and with electrons. At high velocities, as soon as the energy of collision exceeds the lowest excitation potential, the collisions become inelastic, and this leads to additional dispersion of the energies of the colliding bodies. The stationary nature and discreteness of the energy levels of atoms, as of systems of microparticles in general (e.g. molecules, crystals, atomic nuclei) are due to the fact that they obey quantum laws.

The Stratified Structure of the Electron Cloud

The arrangement of the energy levels of the atom shows that its electron cloud has a stratified structure. This can be seen from such observed facts as: (*a*) the presence of breaks in the series of successive atomic ionisation potentials; (*b*) the presence of discrete breaks in the X-ray absorption spectra; (*c*) the grouping in series of optical and X-ray emission spectra.

The n-th order ionisation potential I_n is equal to the energy needed to tear away an electron from an atom (or, to be more exact, ion) which had previously lost the $(n-1)$th-electron. Table 4 gives the values of the ionisation potentials of various orders.

If the density of the electron cloud were a monotonic function of distance from the nucleus, then the successive ionisation potentials of any atom would also increase monotonically. However, experiment shows that, side by side with the monotonic increase, the transition to the succeeding ionisation potential may in certain cases involve a discontinuous increase. Discontinuities in the ionisation potential are marked in the table by thick lines. They show when electrons are beginning to be removed from a new and deeper shell, in which the energy binding its electrons to the

Table 4

Atomic Ionisation Potentials I_n of Different Orders (Ionisation Potential of the Hydrogen Atom taken as Unity)

Element	Z	Number and type of electrons in the outer shell	I_1	I_2	I_3	I_4	I_5	I_6	I_7	I_8
H	1	s	1·00							
He	2	s^2	1·80	4·0						
Li	3	s	0·4	5·6	9·0					
Be	4	s^2	0·7	1·34	11·3	16				
B	5	s^2p	0·61	1·85	2·8	19·3	25			
C	6	s^2p^2	0·83	1·79	3·52	4·74	29	36		
N	7	s^2p^3	1·07	2·18	3·49	5·43	7·2	40·5	49	
O	8	s^2p^4	1·0	2·58	4·04	5·69	8·1	10·1	54·5	64 *K*-shell
F	9	s^2p^5	1·33	2·57	4·62	6·40	7·56	11	13·6	70 2 electrons
Ne	10	s^2p^6	1·58	3·02	4·67					
Na	11	s	0·38	3·5	5·24					
Mg	12	s^2	0·56	1·10	5·89			*L*-shell		
Al	13	s^2p	0·44	1·38	2·09	9·0	8	electrons		
Si	14	s^2p^2	0·60	1·20	2·46	3·32	12·47			
P	15	s^2p^3	0·81	1·46	2·22	3·53	4·78			
S	16	s^2p^4	0·76	1·72	2·58	3·48	4·93	6·47		
Cl	17	s^2p^5	0·96	1·73	2·93	3·50	5·00	6·8	8·0	
A	18	s^2p^6	1·16	2·05	3·01	12·6				
K	19	s	0·32	2·34	3·47					
Ca	20	s^2	0·45	0·87	3·76			*M*-shell		
Sc	21	s^2d	0·49	0·95	1·82	5·33	8	electrons		
Ti	22	s^2d^2	0·50	1·00	2·03	3·18	7·07			
V	23	s^2d^3	0·5	1·04	1·95	3·59	5·05	9·05		

nucleus is greater than in the previous, outer shell. It can be seen from Table 4 that as the number of electrons is increased, inner electron shells K, L and M are formed. As each of these is filled up, one of the principal periods in the periodic system is completed. The capacities of these shells are respectively 2, 8 and 8 electrons.

The observation that the *optical* spectra of atoms form series shows that the *unfilled* excited energy levels of atoms are arranged in strata, while the same observation for X-ray spectra shows that the *filled* electron levels of atoms are stratified. Moseley's Law describes the analogy between the X-ray spectra of different atoms and establishes the monotonic dependence of the energy levels on the nuclear charge.

This is an indication that the inner electron shells of different elements all have the same structure.

Combining the data from optical and X-ray spectra shows that atoms possess a single discrete system of energy levels. Depending on the number of electrons in the atom, a part of the lower levels is filled up, but the remaining levels stay unfilled and can be used for intra-atomic electron transitions when the atoms are excited. The system of energy levels in multi-electron atoms is connected with the system of energy levels of the hydrogen atom, and is found by solving the problem of the motion of an electron in the Coulomb field of the nucleus.

3. The Hydrogen Atom

Separation of the Translational and Rotational Degrees of Freedom

The hydrogen atom is the simplest atom; it consists of a nucleus of mass m_1, the charge on which, for the sake of generality, we shall take as $+Ze$, and an electron of mass m_2 and charge $-e$. The electron and the nucleus rotate about their common centre of gravity; in addition, the atom as a whole can undergo translational motion. Let (x_1, y_1, z_1) be the coordinates of the nucleus, and $(x_2, y_2\ z_2)$ those of the electron. The distance between the electron and the nucleus is

$$r = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2}. \qquad \ldots [2.6]$$

The potential energy of the atom in the absence of an external field is

$$U(r) = -\frac{Ze^2}{r}. \qquad \ldots [2.7]$$

The Hamiltonian for the hydrogen atom is

$$H = \tfrac{1}{2}m_1(\dot{x}_1^2+\dot{y}_1^2+\dot{z}_1^2)+\tfrac{1}{2}m_2(\dot{x}_2^2+\dot{y}_2^2+\dot{z}_2^2)+U(r). \qquad \ldots [2.8]$$

We introduce polar coordinates, placing their origin at the centre of gravity of the atom, Cartesian coordinates of which are (x, y, z). Let θ and ϕ be the polar angles, and a_1 and a_2 the radial distances of the nucleus and the electron:

$$a_1 = -\frac{m_1}{m_1+m_2}r, \qquad a_2 = \frac{m_2}{m_1+m_2}r. \qquad \ldots [2.9]$$

The Cartesian coordinates of the proton and the electron are connected with the polar coordinates by the relationships

$$x_i = x\mp\frac{\mu}{m_i}r\sin\theta\cos\phi; \quad y_i = y\mp\frac{\mu}{m_i}r\sin\theta\sin\phi; \qquad \ldots [2.10]$$

$$z_i = z\mp\frac{\mu}{m_i}r\cos\theta,$$

where the minus sign corresponds to the proton ($i = 1$), and the plus sign to the electron ($i = 2$), and μ is the reduced mass of the atom:

$$\mu = \frac{m_1 m_2}{m_1+m_2}. \qquad \ldots [2.11]$$

In polar coordinates the Hamiltonian function [*8*] takes the form

$$H = \tfrac{1}{2}(m_1+m_2)[\dot{x}^2+\dot{y}^2+\dot{z}^2]+\tfrac{1}{2}\mu[\dot{r}^2+r^2\dot{\theta}^2+(r^2\sin^2\theta)\dot{\phi}^2]-\frac{Ze^2}{r}. \qquad \ldots [2.12]$$

Here the first term is the translational kinetic energy, and the second the rotational kinetic energy, of the atom. Changing to the Hamiltonian operator in accordance with the general rule, we obtain

$$\boldsymbol{H} = -\frac{h^2}{8\pi^2(m_1+m_2)}\nabla^2_{xyz}-\frac{h^2}{8\pi^2\mu}\nabla^2_{r\theta\phi}-\frac{Ze^2}{r}. \qquad \ldots [2.13]$$

Schrödinger's equation for the hydrogen atom has the form

$$\frac{h^2}{8\pi^2(m_1+m_2)}\nabla^2_{xyz}\Psi+\frac{h^2}{8\pi^2\mu}\nabla^2_{r\theta\phi}\Psi+\left(E'+\frac{Ze^2}{r}\right)\Psi = 0, \qquad \ldots [2.14]$$

where E' is the total energy of the atom. The independence of the translational and rotational motions allows us to write the wave function in the form

$$\Psi(x,y,z;r,\theta,\phi) = X(x,y,z)\psi(r,\theta,\phi). \qquad \ldots [2.15]$$

Substitution of [*15*] into [*14*] produces two equations:

$$\nabla^2_{xyz}X+\frac{8\pi^2(m_1+m_2)}{h^2}(E'-E)X = 0, \qquad \ldots [2.16]$$

$$\nabla^2_{r\theta\phi}\psi+\frac{8\pi^2\mu}{h^2}\left(E+\frac{Ze^2}{r}\right)\psi = 0. \qquad \ldots [2.17]$$

Equation [*17*] is the same as equation [*1.153*] for a centrally symmetric system and describes the rotational motion of the atom; here E is the internal energy of the atom. Equation [*16*] is the equation for a free particle of mass m_1+m_2 and kinetic energy $E-E'$ and describes the translational motion of the atom.

The Equation for the Radial Part of the Wave-Function

In equation [*1.185*] we now introduce the energy of the Coulomb interaction, and for simplicity we shall write the equation in dimensionless length and energy units:

$$q = \frac{r}{a} \quad \text{and} \quad \varepsilon = \frac{E}{E_1}, \qquad \ldots [2.18]$$

where a is the radius of the first Bohr orbit, and E_1 is the ionisation energy of the hydrogen atom. Thus

$$a = \frac{h^2}{4\pi^2\mu e^2} = 0{\cdot}529\times10^{-8}\text{ cm.}; \quad E_1 = \frac{2\pi^2\mu e^4}{h^2} = \frac{e^2}{2a} = 13{\cdot}56\text{ e}V, \qquad \ldots [2.19]$$

Doing this we obtain

$$\frac{d^2\rho}{dq^2}+\left[\varepsilon+\frac{2Z}{q}-\frac{l(l+1)}{q^2}\right]\rho = 0. \qquad \ldots [2.20]$$

It follows from [*1.192*] and [*1.197*] that the asymptotic solutions of this equation near the centre and at a great distance from the centre (for $E<0$) have the form

$$\rho_1(q) = A_1 q^{l+1}, \quad \rho_2(q) = A_2 e^{-\alpha q}, \qquad \ldots [2.21]$$

where

$$\alpha = +\sqrt{-\varepsilon}. \qquad \ldots [2.22]$$

The solutions [*21*] must be 'fitted together' into a single function. We shall do this by means of a power series and shall suppose that $\rho(q)$ can be represented in the form

$$\rho(q) = q^{l+1}(a_0 + a_1 q + a_2 q^2 + \ldots)e^{-\alpha q} = e^{-\alpha q}\sum_{\nu=0}^{\infty} a_\nu q^{\nu+l+1}, \qquad \ldots [2.23]$$

where $a_0, a_1, a_2, \ldots$ are constant coefficients. The first factor in [*23*] must give the behaviour of the function $\rho(q)$ near the nucleus, and the exponential factor must express the decay of the function at great distances from the nucleus. We shall find the values of the coefficients a_ν for which [*23*] is a solution of equation [*20*]. Differentiation of [*23*] twice gives

$$\frac{d^2\rho}{dq^2} = \alpha^2 e^{-\alpha q}\sum_{\nu=0}^{\infty} a_\nu q^{\nu+l+1} - 2\alpha e^{-\alpha q}\sum_{\nu=0}^{\infty} a_\nu(\nu+l+1)q^{\nu+l} +$$

$$+ e^{-\alpha q}\sum_{\nu=0}^{\infty} a_\nu(\nu+l+1)(\nu+l)q^{\nu+l-1}. \qquad \ldots [2.24]$$

The first term on the R.H.S. is equal to $-\varepsilon\rho$; it cancels out with the same term in equation [*20*]. Contracting the remaining terms into the general factor $e^{-\alpha q}$, and grouping together terms of the same degree, we obtain, after substituting [*23*] and [*24*] into [*20*],

$$\sum_{\nu=0}^{\infty}\{[(\nu+l+2)(\nu+l+1) - l(l+1)]a_{\nu+1} - [2\alpha(\nu+l+1) - 2Z]a_\nu\}q^{\nu+l} = 0. \qquad \ldots [2.25]$$

The equation thus obtained must hold for all values of q from 0 to ∞. This is possible only if the coefficient of each power of q vanishes. Hence we obtain a recurrence relationship between the coefficients a_n and a_{n+1}

$$a_{\nu+1} = \frac{2\alpha(\nu+l+1) - 2Z}{(\nu+l+2)(\nu+l+1) - l(l+1)}a_\nu, \quad \nu = 0, 1, 2\ldots, \qquad \ldots [2.26]$$

which enables us to find all the coefficients of series [*23*] in terms of the first.

For large values of ν in [*26*] we can neglect Z and l, and so obtain

$$a_{\nu+1} = 2\alpha\frac{a_\nu}{\nu} \quad (\nu \to \infty). \qquad \ldots [2.27]$$

Formula [*27*] gives the coefficients in the series expansion of an exponential function. So approximately, for large values of q, we obtain

$$\Sigma a_\nu q^\nu \approx e^{2\alpha q}. \qquad \ldots [2.28]$$

Series [*28*] tends to infinity when $q \to \infty$, and this divergence is not removed by the factor $e^{-\alpha q}$ which enters into the expression for $\rho(q)$. Divergence of $\rho(q)$ can be avoided if the infinite series [*23*] can be replaced by a finite series, i.e. a polynomial. For this, we break off the series at $v = n$ by arranging that the coefficient

$$a_{v+1} = a_{n_r+1} = 0. \qquad \ldots [2.29]$$

From the recurrence relation [*28*] it follows that all the succeeding coefficients of the expansion will also vanish. As can be seen from [*26*], condition [*29*] is observed if

$$\alpha = \frac{Z}{n_r+l+1} = \frac{Z}{n}, \qquad \ldots [2.30]$$

where

$$n = n_r+l+1; \quad n_r = 0, 1, 2 \ldots \quad n = 0, 1, 2 \ldots \qquad \ldots [2.31]$$

Formula [*30*] gives the energy quantisation condition. Substituting for α from [*22*] and changing to the usual energy units, we have

$$E = -\frac{2\pi^2 \mu Z^2 e^4}{h^2} \cdot \frac{1}{n^2}. \qquad \ldots [2.32]$$

Energy Levels of the Hydrogen Atom

The diagram of the energy levels of hydrogen and hydrogen-like ions, drawn in accordance with formula [*32*], (Fig. 26), shows clearly the steady narrowing of the gaps between the energy levels with increasing principal quantum numbers. Transitions between these later levels correspond to low excitation energies. When the energy levels are perturbed, for example, as a result of interaction between atoms, neighbouring levels may easily overlap. Spontaneous transitions of an electron from excited levels to lower levels are accompanied by radiation of a line spectrum. The frequencies of the lines belonging to the various series for the hydrogen atom (without fine structure details) are shown in the diagram in $\mathrm{cm.}^{-1}$ The spectroscopic data fully confirm the correctness of this diagram. The ground state energy level, corresponding to the lowest value of the principal quantum number, $n = 1$, determines the ionisation energy of the hydrogen atom. $I = E_1 = 13{\cdot}56$ eV. This value for the ionisation energy agrees well with experiment.

Quantum Numbers

The equation for the radial part is soluble for integral values of the quantum number n [*31*]. The equation for the angular part, as was shown in Ch. 1, §7, is soluble for values of the quantum numbers l and m given by formula [*1.171*]. Corresponding to the three degrees of freedom of the electron there are the three quantum numbers:

$$\left.\begin{array}{ll} \text{principal:} & n = 1, 2, 3, \ldots \\ \text{orbital:} & l = 0, 1, \ldots, n-1 \\ \text{magnetic:} & m = 0, \pm 1, \pm 2, \ldots, \pm l \end{array}\right\} \qquad \ldots [2.33]$$

The principal quantum number determines the energy E of the atom. This energy does not depend on the quantum numbers l and m; consequently, the energy levels of the hydrogen atom are degenerate. When the degeneracy is removed, as it is in multi-electron atoms because of disturbance of the spherical symmetry of the field, the quantum number l determines the sub-levels into which the energy levels of the hydrogen atom split. To denote these sub-levels letter symbols are used, borrowed from spectroscopy:

orbital quantum number: $l = 0,\ 1,\ 2,\ 3.$
spectroscopic symbol: $s\ \ p\ \ d\ \ f.$

The orbital quantum number determines the mechanical M_e and the magnetic M_m orbital moments

$$M_e = \frac{h}{2\pi}\sqrt{l(l+1)}, \quad M_m = \sqrt{l(l+1)}\,M_B, \qquad \ldots [2.34]$$

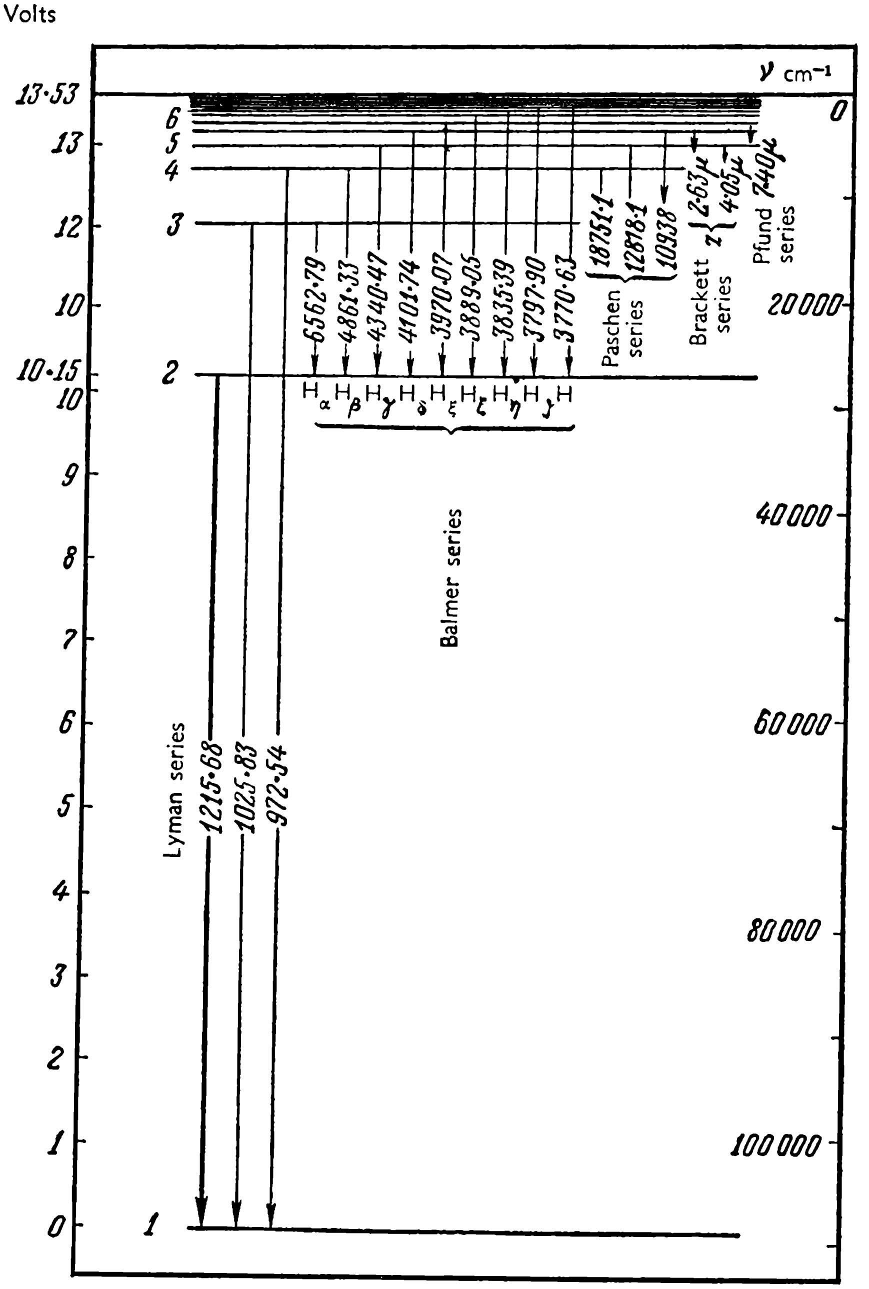


FIG. 26. Diagram of levels and spectrum of the hydrogen atom.

where M_B is the Bohr magneton. The gyromagnetic ratio of the orbital moments:

$$\frac{M_m}{M_e} = \frac{e}{2\mu c} \qquad \ldots [2.35]$$

is half the ratio of the spin moments [2]. In a magnetic field of intensity H the atom acquires additional potential energy

$$U_M = -HM_m \cos\theta, \qquad \ldots [2.36]$$

where θ is the angle between the direction of the magnetic moment and the direction of the field. This energy is small by comparison with the energy of electrostatic interaction between an electron and the nucleus, but it is nevertheless sufficient to remove the degeneracy and split the levels. The magnetic field causes precession of the magnetic moment vector and of the angular momentum vector, which is parallel to it, about the direction of the magnetic field, the angle θ remaining constant. In agreement with the space quantisation rule expressed by formula [*1.181*], only strictly defined values of the angles of orientation θ_m are possible, each corresponding to an integral value of the projection of the momentum on the direction of the field:

$$M_z = \frac{h}{2\pi} m; \qquad m = 0, \pm 1, \pm 2, \ldots, \pm l. \qquad \ldots [2.37]$$

To the possible values of the magnetic quantum number m there correspond $(2l+1)$ orientations of the moment M_e in the magnetic field.

The quantum numbers n, l, m define the scheme of possible energy levels of electrons in the atom. These levels are divided into groups, depending on the principal quantum number. The scheme of energy levels of a one-electron atom is shown in Fig. 27 in the form of ordered quantum cells. Each group of levels consists of subgroups s, p, d. For the one-electron atom the energies of different levels within each group are the same. The number of cells in the group n is

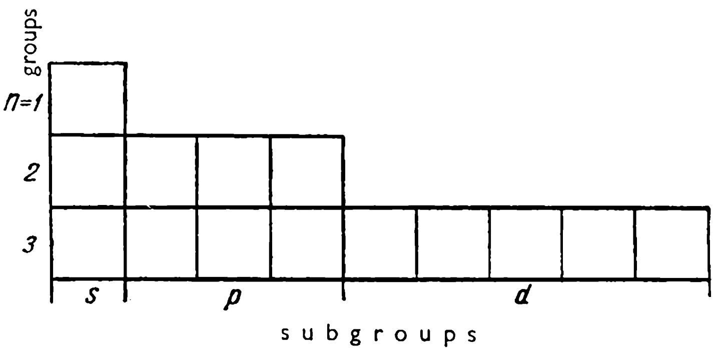


FIG. 27. Scheme of quantum cells for the one-electron atom.

$$\sum_{l=0}^{n-1} (2l+1) = n^2. \qquad \ldots [2.38]$$

Each cell corresponds, in accordance with Pauli's principle, to two electron energy states, representing the two different spin orientations. Each state is now defined by four quantum numbers: n, l, m, s. The total number of states in the group n is $2n^2$. We obtain the following capacities for the successive electron groups:

$$2, 8, 18, 32, \ldots, 2n^2. \qquad \ldots [2.39]$$

These numbers are equal to the lengths of the periods in the periodic table, but the sequence is not the same.

The Radial Function of the Hydrogen Atom. The Laguerre Polynomial

Let us substitute the coefficients [*26*], for the value of α given by [*30*], into [*23*]; then we find

$$\rho(q) = a_0 e^{-\frac{Z}{n}q} q^{l+1} \left\{1 - \frac{n-l-1}{1!(2l+2)}\left(\frac{2Zq}{n}\right) + \frac{(n-l-1)(n-l-2)}{2!(2l+2)(2l+3)}\left(\frac{2Zq}{n}\right)^2 \right.$$
$$\left. + \ldots - \frac{(n-l-1)(n-l-2)\ldots 1}{(n-1)!(2l+2)(2l+3)\ldots(2l+n_r)}\left(\frac{2Zq}{n}\right)^{n_r}\right\}. \qquad \ldots [2.40]$$

Introducing the new variable

$$\xi = \frac{2Zq}{n} = \frac{2Zr}{na} \qquad \ldots [2.41]$$

and substituting [*40*] into [*1.184*], we have the radial wave function of the hydrogen atom

$$R_{nl}(\xi) = N_{nl} e^{-\frac{1}{2}\xi} \xi^l L_{n+l}^{2l+1}(\xi), \qquad \ldots [2.42]$$

where N_{nl} is a normalising factor whose value is given by the normalisation condition

$$\int_0^\infty R_{nl}^2 r^2 dr = 1,$$

and L_{n+l}^{2l+1} is the polynomial in braces in [*40*]. This polynomial can be obtained by differentiating,

$$L_k^s(\xi) = \frac{d^s}{d\xi^s} L_k(\xi) \qquad \ldots [2.43]$$

where $L_k(\xi)$ is the Laguerre polynomial:

$$L_k(\xi) = e^\xi \frac{d^k}{d\xi^k}(e^{-\xi}\xi^k). \qquad \ldots [2.44]$$

The Wave-functions of the Hydrogen Atom

The complete wave function of the hydrogen atom is equal to the product of the radial function [*42*] and the spherical functions [*1.176*]:

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) = R_{nl}(r)\Theta_{lm}(\theta)\Phi_m(\phi).$$

Table 5 gives the radial function for certain values of the quantum numbers n and l.

Table 5

Normalised Radial Wave Functions for Hydrogen-like Atoms

$$\left(\rho = \frac{Z}{a_0} r;\ a_0 = \frac{h^2}{4\pi^2 \mu e^2} = 0{\cdot}529 \times 10^{-8}\ \text{cm.—radius of the first Bohr orbit}\right)$$

n	l	$R_{nl}(r)$
1	0	$R_{10} = 2\left(\frac{Z}{a_0}\right)^{3/2} e^{-\rho}$
2	0	$R_{20} = \frac{1}{2\sqrt{2}}\left(\frac{Z}{a_0}\right)^{3/2} (2-\rho)e^{-\rho/2}$
	1	$R_{21} = \frac{1}{2\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} \rho e^{-\rho/2}$
3	0	$R_{30} = \frac{2}{81\sqrt{3}}\left(\frac{Z}{a_0}\right)^{3/2} (27-18\rho+2\rho^2)e^{-\rho/3}$
	1	$R_{31} = \frac{4}{81\sqrt{6}}\left(\frac{Z}{a_0}\right)^{3/2} (6\rho-\rho^2)e^{-\rho/3}$
	2	$R_{32} = \frac{4}{81\sqrt{30}}\left(\frac{Z}{a_0}\right)^{3/2} \rho^2 e^{-\rho/3}$

Table 6 gives the values of the angular part of the eigenfunction $\Theta_{lm}(\theta)$.

Table 6

Normalised Angular Wave Functions $\Theta_{lm}(\theta)$

l	m	$\Theta_{lm}(\theta)$	l	m	$\Theta_{lm}(\theta)$
0	0	$\Theta_{0,0} = \sqrt{\frac{1}{2}}$	3	0	$\Theta_{3,0} = \sqrt{\frac{63}{8}}\left(\frac{5}{8}\cos^3\theta - \cos\theta\right)$
1	0	$\Theta_{1,0} = \sqrt{\frac{3}{2}}\cos\theta$		± 1	$\Theta_{3,\pm 1} = \sqrt{\frac{21}{32}}(5\cos^2\theta - 1)\sin\theta$
	± 1	$\Theta_{1,\pm 1} = \sqrt{\frac{3}{4}}\sin\theta$		± 2	$\Theta_{3,\pm 2} = \sqrt{\frac{105}{16}}\sin^2\theta\cos\theta$
2	0	$\Theta_{2,0} = \sqrt{\frac{5}{8}}(3\cos^2\theta - 1)$		± 3	$\Theta_{3,\pm 3} = \sqrt{\frac{35}{32}}\sin^3\theta$
	± 1	$\Theta_{2,\pm 1} = \sqrt{\frac{15}{4}}\sin\theta\cos\theta$			
	± 2	$\Theta_{2,\pm 2} = \sqrt{\frac{15}{16}}\sin^2\theta$			

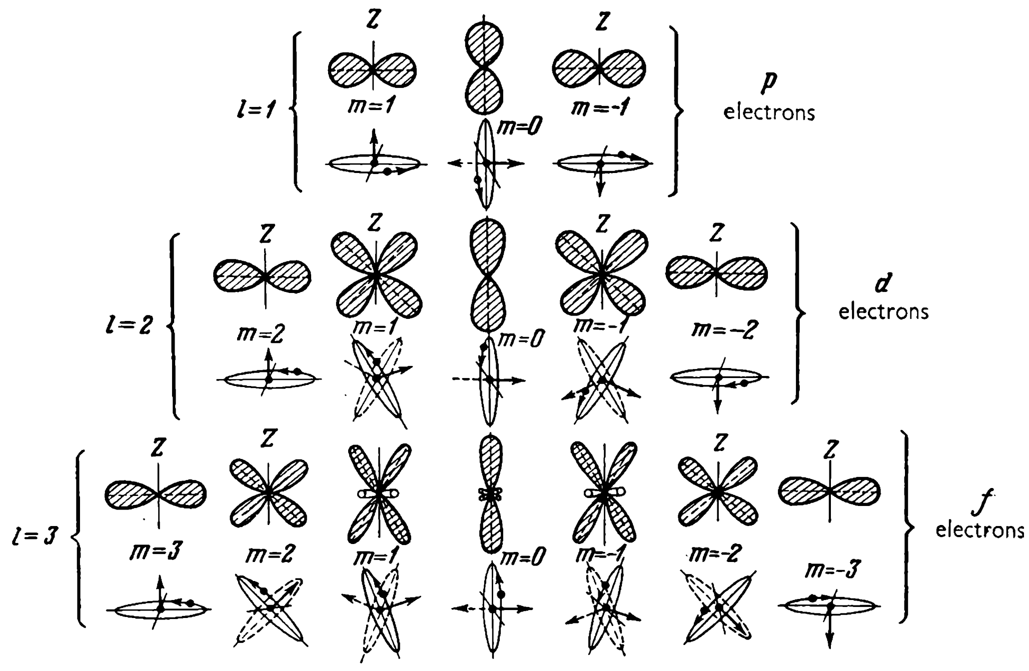


FIG. 28. Angular electron density distribution in the hydrogen atom for various states *l*, *m*.

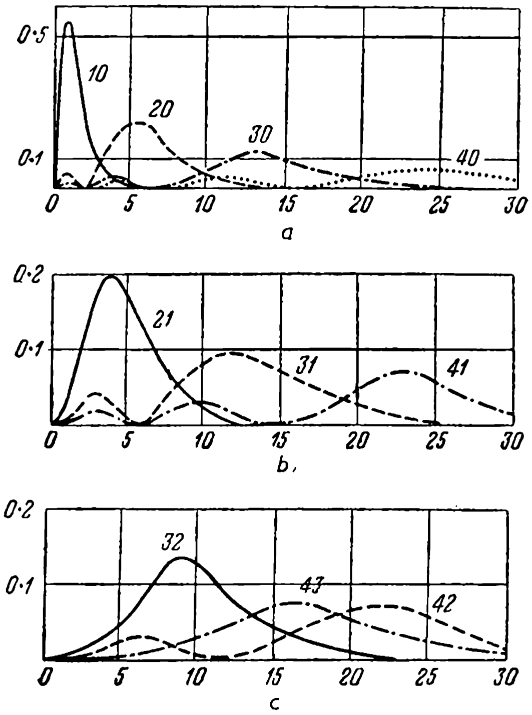


FIG. 29. Radial electron density distribution in the hydrogen atom for various states *l*, *m*: (*a*) *s*-state ($l=0$), (*b*) *p*-state ($l=1$), (*c*) *d*-state ($l=2$), and 4*f*-state ($l=3$).

Electron Density Distribution for Different States of the Hydrogen Atom

The probability of finding an electron which is in the quantum state n, l, m in the neighbourhood of the point r, θ, ϕ, in the element of volume $dr = r^2 dr \sin\theta d\theta d\phi$ is given by the expression

$$w_{nlm}^2(r, \theta, \phi)\, dv = |\psi_{nlm}(r, \theta, \psi)|^2 dv = R_{nl}^2(r)\, \Theta_{lm}^2(\theta)\, dv. \qquad \ldots [2.45]$$

This probability does not depend on the angle ϕ so the electron density distribution has the symmetry of a solid of revolution about the Z-axis, with respect to which the projection of the moment of momentum is quantised.

If we integrate function [*45*] with respect to r from 0 to ∞, we obtain the probability of finding the electron in the solid angle $d\Omega$:

$$w_{lm}(\theta)\, d\Omega = \Theta_{lm}^2(\theta) \sin\theta\, d\theta\, d\phi. \qquad \ldots [2.46]$$

Fig. 28 shows the angular electron density distribution for various states l, m. The radial distances are proportional to the corresponding values of the probability w_{lm}. These graphs show that the s-states of the atom have spherical symmetry. The p, d, f etc. states are characterised by a marked concentration of electron density along certain straight lines and planes. For comparison Fig. 28 gives the positions of the electron orbits according to the Bohr theory. The maxima of the electron density are in the directions of the Bohr orbits.

If the function [*45*] is integrated with respect to both the angles, we obtain the probability of finding the electron in a spherical shell lying between radii r and $r+dr$:

$$w_{nl}(r)\, dr = R_{nl}^2(r)\, r^2 dr. \qquad \ldots [2.47]$$

Fig. 29 gives these probabilities for various states l, m. For higher quantum states the electron density varies concentrically. In these cases we can distinguish a spherical shell in which the probability of finding the electron is greatest, but the possibility is not excluded of some 'leakage' of the electron into regions, nearer to the nucleus, where the electron is found less of the time. These electron states are represented in the Bohr model by elliptic orbits. A 'leakage' region corresponds to passage of the electron across a perihelion. The undulatory variation of the radial probability w_{nl} shows that it is impossible to establish complete agreement with the Bohr orbits.

4. The Multi-electron Atom

Schrödinger's Equation for a System of Particles. Approximate Methods of Solution

For the multi-electron atom we can write down the appropriate Schrödinger equation of the type [*1.38*], but owing to mathematical difficulties it is impossible to solve it exactly. Therefore approximate methods of solution are employed. One simple method, suggested by Thomas and Fermi, arises from the statistical model of the atom and is applicable to atoms containing a sufficiently large number of electrons (starting, say, from the middle of the Periodic Table). By this method we can determine the mean radial electron density distribution, which is essential for finding the atomic scattering function in X-ray and electron diffraction. The analogous problem for light atoms can be solved by the method of the 'self-consistent' field, proposed by

Hartree and developed by Fock. This method is based on a series of successive approximations, and requires a vast amount of laborious computation.

The complexity of the problem for a multi-electron atom is due to the fact that each of the electrons in this case interacts not only with the nucleus but also with the other electrons. The resultant field acting on the electrons varies according to a more complicated law than Coulomb's. But for the most part ($\sim 70\%$) the potential energy of the electrons is determined by their interaction with the nucleus, because of the size of the charge on the latter. So the resulting field can be looked on as quasi-central. As in the one-electron atom, the state of an electron is defined by four quantum numbers, n, l, m, s, having the same physical meaning as before. The energy level scheme of the hydrogen atom can be taken as the basis. However the sequence of levels is different because the lower energy states are fully occupied by electrons.

> In the other many-particle problem of quantum mechanics (namely that of the atomic nucleus, in which there is no central particle), the interaction is approximately the same for all the particles. So the scheme of nuclear levels is fundamentally different from that of the atom. The sequence of levels in the nucleus is: 1*s*, 2*p*, 3*d*, 2*s*, 4*f*, 3*p*, 5*g* etc. At the same time for some molecules we can keep the terminology of a single atom, which could be considered as formed from the molecule by coalescence of the nuclei.

Scheme of Energy Levels in the Multi-electron Atom

In the multi-electron atom the energy of an electron depends not only on the principal but also on the orbital quantum number, and so the hydrogen spectrum levels are split up. As the quantum number n increases, the distances between neighbouring levels in the hydrogen atom lessen. This splitting can lead to overlapping of the levels belonging to different values of n, that is, to deviation from the energy level sequence of the hydrogen atom. The reason for this is the screening effect of the electrons in the core. Let us consider from this point of view the 3*d*- and 4*s*-states. It can be seen from Fig. 29 that the maximum of the 3*d*-state is nearer to the nucleus than the principal maximum of the 4*s*-state. So in the hydrogen atom the 3*d*-level lies lower than the 4*s*-level. In the multi-electron atom, starting with the atom of argon, there is a core of 18 internal electrons, forming a spherically symmetrical group. The core electrons screen the nucleus and lessen its effective charge. So the energy of the 3*d*-state is considerably less. An electron in the 4*s*-state spends a certain amount of time near the nucleus, that is, inside the core and during this time it hardly experiences the screening effect of the latter at all. For this reason the energy of the 4*s*-state in the multi-electron atom is diminished to a lesser degree, and is greater than the energy of the 3*d*-state. In the Bohr model the 4*s*-state corresponds, as seen above, to an elliptical orbit with a high eccentricity, 'diving' deeply into the interior of the core and passing close to the nucleus at its perihelion. The 3*d*-state corresponds to a circular orbit passing wholly outside the core.

The sequence of energy levels in the multi-electron atom is shown in Fig. 30 in the form of a scheme of quantum cells. This sequence is confirmed by a number of experimental data, such as the spectra of multi-electron atoms, the sequence of period lengths in the Periodic Table etc., and agrees with the approximate quantum mechanical calculation. The Thomas-Fermi method can be used to find the values of the atomic number Z for which the p-, d-, f-, etc. electrons first appear. The theory gives the following formula

$$Z = 0{\cdot}155\,(2l+1)^3,$$

which has been verified to within 10%. If instead of 0·155 the coefficient is taken as 0·17, we obtain f $l = 1, 2$ and 3 respectively $Z = 4{\cdot}6$, 21·2 and 58·3. Discarding fractions, we find $Z = 5$, 21 and 58. This result, obtained semi-empirically, agrees with experiment, since p-electrons first appear in the 5th element (boron), d-electrons in the 21st element (scandium) and f-electrons in the 58th (cerium). Extrapolating the formula for $l = 4$, we find that g-electrons should first appear in an as yet unknown 124th element.

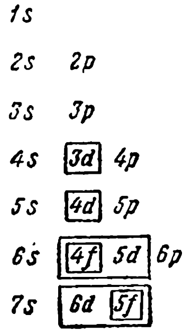


FIG. 30. Scheme of quantum cells of the multi-electron atom.

The start of a new period, the n-th, in the Periodic Table is connected with a new electron shell beginning to be built up, that is, with the appearance of an ns-electron. Hence (Fig. 30) we obtain the figures for the periods: 2, 8, 8, 18, 18, 32, . . . , in agreement with Mendeleev's law.

Order of Filling the Electron Subgroups (Hund's Rules)

The order of filling up the quantum cells in separate electron subgroups as the charge on the nucleus and the number of electrons in the atoms increase is given by Hund's rules, which were discovered by analysis of atomic spectra. Great importance attaches to the total orbital momentum L and the total spin momentum S of the multi-electron atom. Added together, they form the resultant momentum J.

In forming the total spin or orbital momentum, the component momenta are so orientated to one another as to maximise the magnitude of the resultant S *or* L; (for spins this means that the *parallel* orientation is preferred).

In forming the resultant momentum J, *when the first half of the states in the electron subgroups is filled up, the component momenta* L *and* S *are parallel but opposite in sense,* $J = L-S$, *but when the second half of the states is filled up they are parallel and in the same sense* (*the latter is a consequence of the Pauli principle*).

Fig. 31 shows the successive filling up of the cells of a p-subgroup with electrons. Each arrow corresponds to one electron and indicates the direction of spin. A similar order is observed in the filling up of the d- and f-electron subgroups. Fig. 32 shows the total momenta S, L and J corresponding to different degrees to which the p-, d- and f-subgroups are filled. For completely filled subgroups the total momenta are zero. So the curves of Fig. 32 also represent the total momenta of atoms and ions with the corresponding electron configurations in the inner core. The curves of Fig. 32*a* give the properties of atoms in the second period from Li to Ne, those of Fig. 32*b* the properties of bivalent ions in the fourth period from Ca to Zn and the curves of Fig. 32*c* the trivalent ions from La to Lu.

The electron charge distribution in filled subgroups is characterised by spherical symmetry. In a filled electron subgroup with quantum number l there are $2\,(2l+1)$

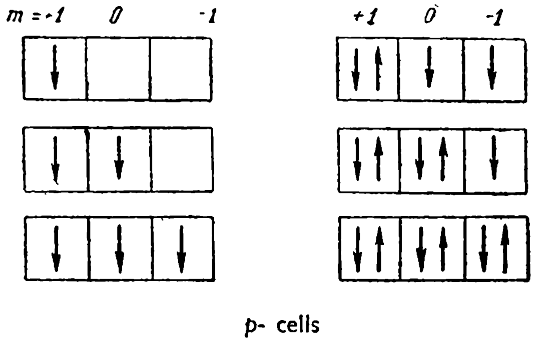


FIG. 31. Order of filling-up p-cells with electrons.

electrons with magnetic quantum numbers ranging from $-l$ to $+l$. The electron density distribution for one of the electrons with a definite value m is given according to [*45*], by the formula

$$\rho_m = |\psi|^2 = R_{nl}^2(r)\Theta_{lm}^2(\theta).$$

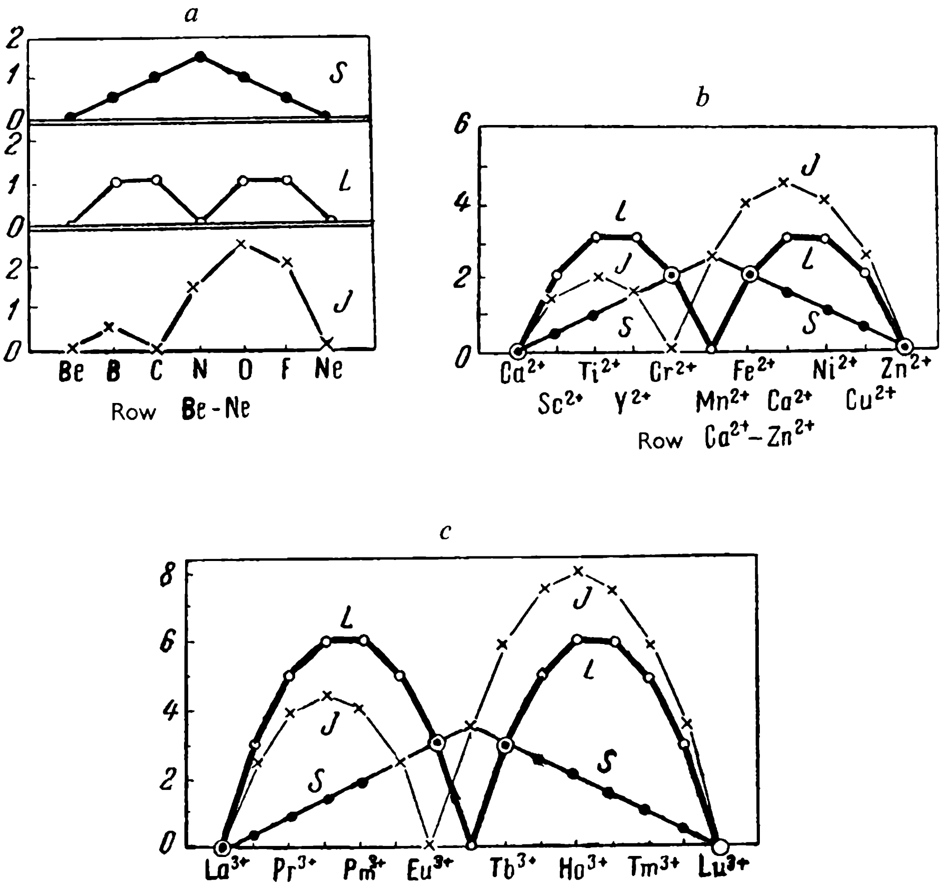


FIG. 32. Total moments S L and J for (*a*) p-, (*b*) d- and (*c*) f- subgroups.

The total electron density for a filled subgroup can be expressed as the sum:

$$\sum_{m=-l}^{+l} \rho_m = R_{nl}^2(r) \sum_{m=-l}^{+l} \Theta_{lm}^2(\theta) = R_{nl}^2(r) \sum_{m=-l}^{+l} N_\theta^2 [P_l^m(\cos\theta)]^2. \qquad \ldots [2.48]$$

As is to be expected from the properties of spherical functions, their sum in [*48*] does not depend on the angle θ.† Consequently, $\Sigma\rho_m$ does not depend on the spherical angles and varies only with l and n.

5. The Periodic Law and the Periodic System of the Elements

Chemically Simple and Chemically Complex Substances

Matter undergoes various physical and chemical changes and transformations, either spontaneously in nature, or as a result of human agency. Some transformations are connected with changes in the state of aggregation (condensation—sublimation, crystallisation—melting, solid phase transitions, allotropy, polymorphism, etc.). In these transformations the chemical composition does not alter. Other transformations are connected with changes in the chemical composition of the substance or in the valency state of the atoms which form it.

Even before the atomic theory became widely accepted, it was known that substances could be either *simple* or *complex* in their chemical composition. Chemically composite substances can be decomposed into chemically simple substances—*chemical elements.* The latter cannot be decomposed into simpler substances, since they consist of atoms of one kind only. Atoms (or, to be more exact, atomic nuclei, which determine the properties of the atoms) are particles which are indivisible by ordinary chemical methods. In contrast to the known number of chemical compounds, which is vast and virtually inexhaustible,‡ the total number of chemical elements has turned out to be finite and relatively small, of the order of 100, and equal to the number of different kinds of atom.

Stability of the Properties of Elements and their Compounds

From the point of view of classical physics, one of the most striking properties of chemical elements and their compounds is the complete reproducibility of their physical and chemical properties. Despite the mutual interaction of the atoms, the individual properties of a chemical element or compound remain constant and independent of the method by which or substances from which the element or compound has been obtained. This reproducibility, which is both theoretically and practically important, is what Mendeleev meant when he spoke of 'the unchanging essence of eternally changing matter'.

The reproducibility of the physical and chemical properties of a substance when the composition and the thermodynamic parameters which determine the state are kept constant is a consequence of the stationary nature of the energy states and electron configurations of atoms, molecules, crystals, etc. The stationary nature of the states of atoms and microsystems in general is due to the fact that they obey quantum laws. After any perturbation, no matter how violent, even total ionisation (the state of matter at stellar temperatures), provided, of course, that the nuclear charge remains constant, atoms on returning to their ground state, completely recover the configurations of their electron clouds. According to the laws of classical physics, implying continuous changes in properties, atoms of the same element in the unperturbed state would have different electron configurations and would lose their uniqueness. They would become like the systems of celestial mechanics, in which a violent perturbation can lead either to a substantial change or to complete and irreversible disruption. In this event the difference between elements and the very concept of a chemical element would lose any meaning. The existence of chemical elements with constant and unique properties is due to the quantum laws which ensure that *the configuration of the electron cloud of an atom in the ground state is uniquely determined by the charge on the nucleus.*

---

† See e.g. A. Sommerfeld, Wave Mechanics, Dutton; New York, 1929.

‡ At the present time the number of known minerals is ~3,000, of inorganic compounds is ~50,000 and of organic compounds ~2,000,000. The number of metallic compounds and alloys is difficult to determine because of the tendency of metals to form phases of variable composition.

Monotonic and Periodic Properties of the Elements

If the elements are placed in order of increasing atomic number, two groups of properties are discovered: one group changes monotonically; the other periodically (Fig. 33). The monotonic properties include the atomic weight and X-ray spectra. The monotonic variation of atomic weight arises, as we have seen, from the stability condition for atomic nuclei. The monotonic variation of the X-ray spectra, expressed in Moseley's law, is a consequence of the unique structure of the inner electron cores of the atoms and the increase in the energy of inner electrons with increasing nuclear charge. The X-ray spectra of atoms differ very little in the free and bound (i.e. molecules and crystals) states. There is also very little effect on the X-ray spectra caused by the state of aggregation of the substance. This shows that the inner electrons do not play a direct part in interatomic and intermolecular interactions.

Most of the properties of atoms in the free and bound states display regular repetition or, as it is called, periodicity. The length of the period does not remain constant, but increases in a regular manner with increase in the atomic number. The periodic law appears in the optical atomic spectra which arise during electron transitions within the cloud of valence electrons. Similar optical spectra are possessed by, for example, the hydrogen atom and the atoms of alkaline metals: 1H, 3Li, 11Na, 19K, 37Rb, 55Cs and 87Fr. The optical spectra for chemically bound atoms vary considerably. Molecules are characterised by band spectra, and condensed systems (crystals and amorphous or solid bodies) by continuous spectra which in reality are also discrete, but with densely packed levels (quasi-continuous spectrum). This shows that the optical electrons are valence electrons, and play a direct part in chemical bonding.

The periodicity is also clearly shown in the curve of the first ionisation potential for atoms; which is equal to the work done in tearing away the first valence electron (Fig. 34). The maxima of this curve correspond to another group of analogous elements, whose distinguishing feature is their chemical inertness: 2He, 10Ne, 18Ar, 36Kr, 54Xe.

The periodic regularity of chemically *bound atoms* can be traced in a number of properties of elements in the *crystalline state*, for example in the curves of volume (Fig. 33), melting point, coefficient of linear expansion, coefficient of compressibility (Fig. 35). In all cases the distances between turning points of curves showing a

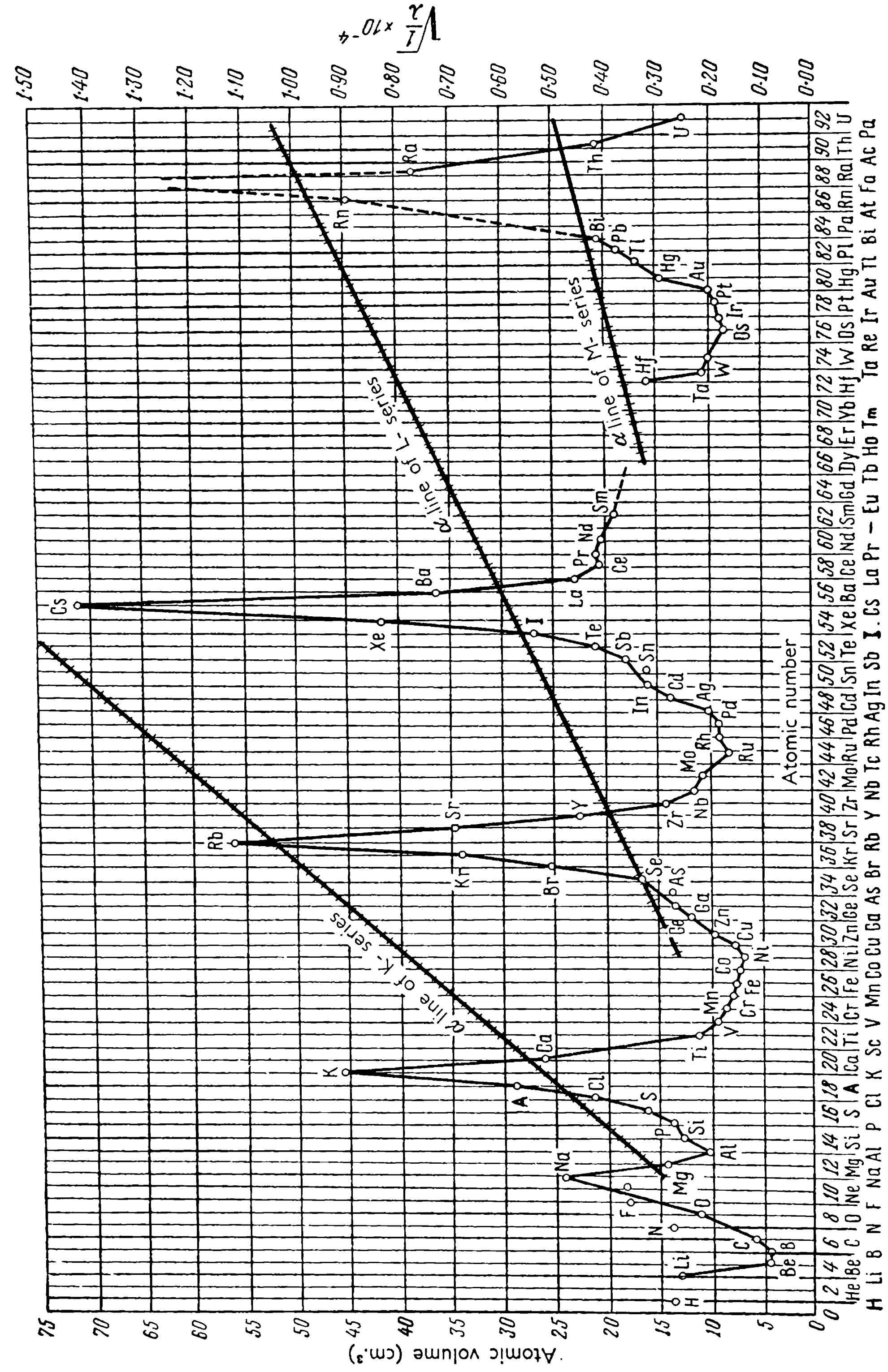


FIG. 33. Dependence of atomic volume of an element on its atomic number.

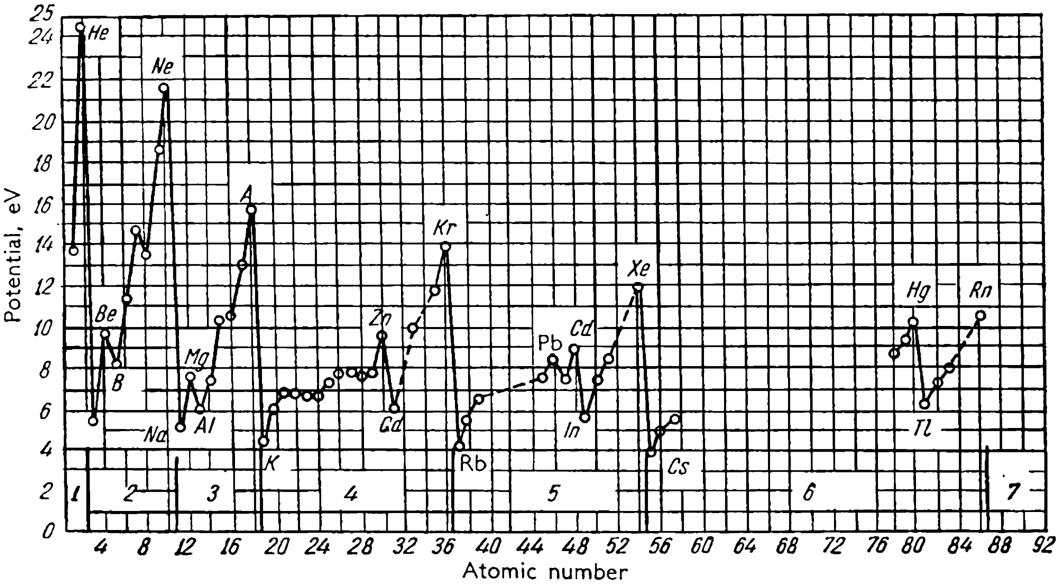


FIG. 34. Dependence of ionisation potential of atoms on their atomic numbers.

particular property against atomic number correspond to the lengths of the periods and always give the same sequence: 2, 8, 18, 18, 32.

The lengths of the periods are equal to the capacities of the electron shells of the atoms, and their sequence coincides with the order of filling the energy levels in the multi-electron atom.

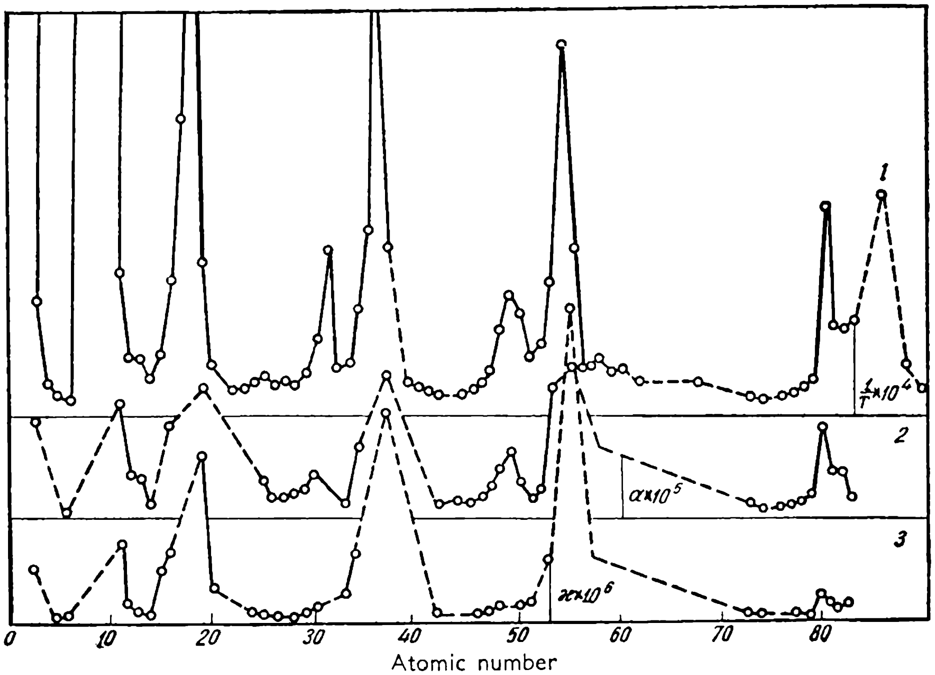


FIG. 35. Periodicity of physical properties: (1) reciprocals of melting temperatures $1/T \times 10^4$; (2) coefficient of linear expansion $\alpha \times 10^5$; (3) coefficient of compressibility $\kappa \times 10^6$ (for the sections of the curves shown dotted, no reliable information exists).

Table 7

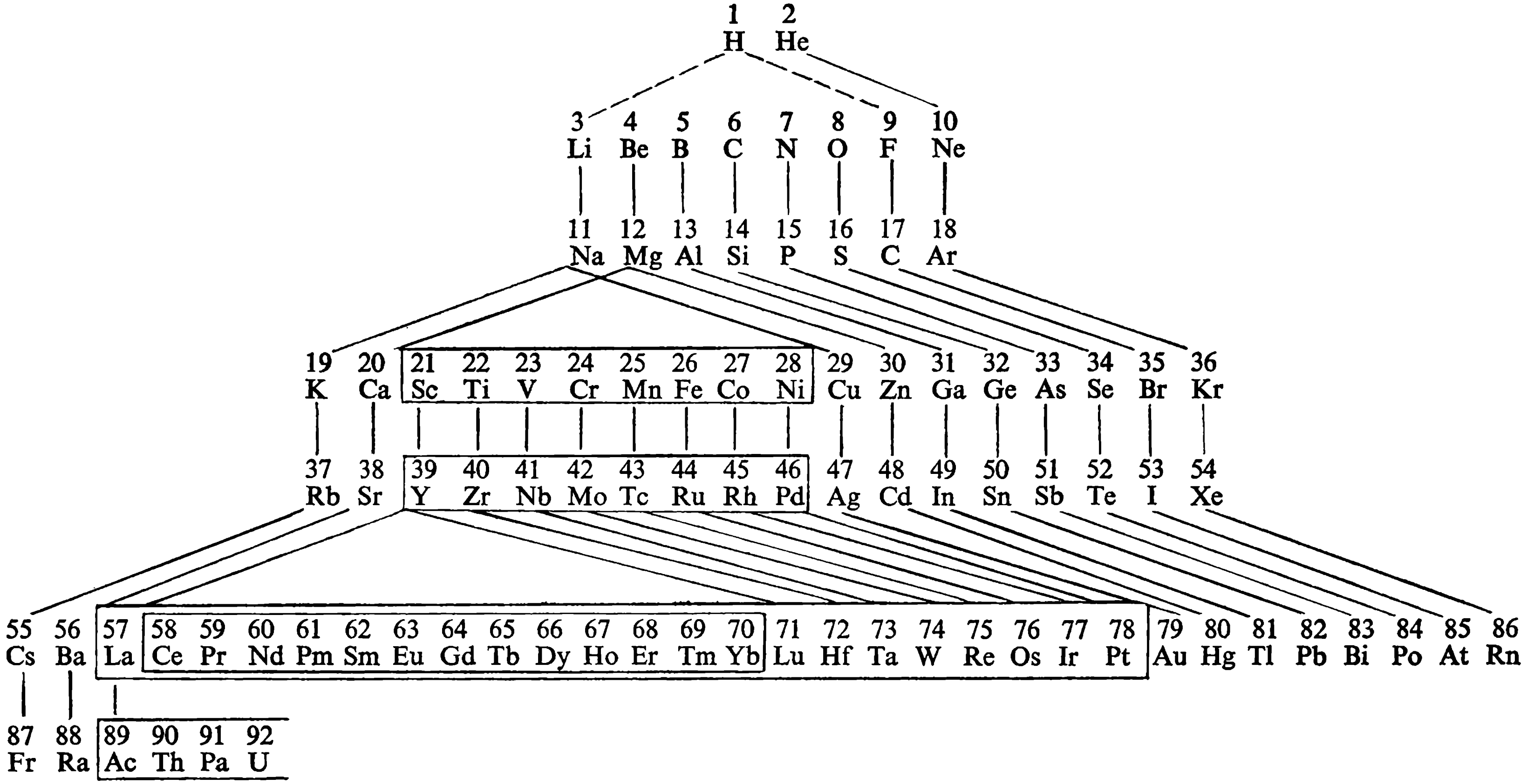
The Periodic System of the Elements

	I	II	III	IV	V	VI	VII	VIII			0
1	1H 1·0080										2He 4·003
2	3Li 6·940	4Be 9·02	5B 10·82	6C 12·01	7N 14·008	8O 16·0000	9F 19·000				10Ne 20·183
3	11Na 22·096	12 Mg 24·32	13Al 26·97	14Si 28·06	15P 30·98	16S 32·06	17Cl 35·457				18Ar 39·944
4	19K 39·096	20Ca 40·08	21Sc 45·10	22Ti 47·90	23V 50·95	24Cr 52·01	25Mn 54·93	26Fe 55·85	27Co 58·94	28Ni 58·69	
	29Cu 63·57	30Zn 65·38	31Ga 69·72	32Ge 72·60	33As 74·91	34Se 78·96	35Br 79·916				36Kr 83·7
5	37Rb 85·48	38Sr 87·63	39Y 88·92	40Zr 91·22	41Nb 92·91	42Mo 95·95	43Tc 99	44Ru 101·7	45Rh 102·91	46Pd 106·71	
	47Ag 107·88	48Cd 112·41	49In 114·76	50Sn 118·70	51Sb 121·76	52Te 127·61	53I 126·92				54Xe 131·3
6	55Cs 132·91	56Ba 137·36	57La† 138·92	72Hf 178·6	73Ta 180·88	74W 183·92	75Re 186·31	76Os 190·2	77Ir 193·1	78Pt 195·23	
	79Au 197·2	80Hg 200·61	81Tl 204·39	82Pb 207·21	83Bi 209·00	84Po 210	85At 211				86Rn 222
7	87Fr 223	88Ra 226·05	89Ac 227·05	‡							

†58–71 lanthanide series	58Ce 140·13	59Pr 140·92	60Nd 144·27	61Pm 145	62Sm 150·43	63Eu 152·0	64Gd 156·9	65Tb 159·2	66Dy 162·46	67Ho 164·94	68Er 167·2	69Tm 169·4	70Yb 173·04	71Lu 174·99

Table 8

The Periodic System (according to Bohr)



Tables I-VII of Appendix 1 at the end of this book show a number of fundamental properties of the elements (ionisation potentials, heats of sublimation, melting points, density in the solid state, electrical conductivity, magnetic properties) which display periodicity. One can understand in the light of this general periodic law a number of particular results and some at first sight unexpected variations ('anomalies') which in the last analysis are due to the structure and properties of atoms.

The Periodic Table of the Elements

The commonest arrangement of the Periodic Table is shown in Table 7.

This table is based on the period 8. The table for the period 18 is given in Tables I-VII (Appendix 1). Both tables are due to D. I. Mendeleev in his *Foundations of Chemistry*. The second originally consisted of 17 groups, since at that time the noble gas group was not known. In the abbreviated form of the table the elements in each group have to be divided into two subgroups A and B, which from the 4th period on, form even and odd series, headed by the alkaline and univalent heavy metals. In group VIII the elements are divided into triads, and, in addition, a ninth (null) group has to be introduced. In the fuller form (Table I, Appendix 1) all the A elements are on the left, and all the B elements on the right. The boundary between the elements of types A and B passes in this case to the left of the column of heavy univalent metals Cu, Ag, Au. In both these forms of the Periodic Table important groups of elements such as the Lanthanides and Actinides lie outside the limits of the table.

Bohr suggested a form of the Periodic Table in the shape of a trapezium (Table 8), which corresponds better with the laws governing the electron structure of atoms. The elements of the B-subgroups lie in the upper and right hand parts of the table. Elements whose atoms are having their *d*-states filled with electrons (transition metals) are enclosed in a rectangle. Elements whose atoms are having their *f*-states filled (rare earth metals—lanthanides, and heavy metals from the actinide group) are enclosed by a double rectangle. Analogous elements are joined by lines.

Metals, Metalloids and Chemically Inert Elements

The elements may be divided according to their ability to form strong interatomic bonds into two unequal groups, namely, the chemically inert and the chemically active elements. The chemically inert elements are the group of noble gases, which consists of 6 elements (He, Ne, Ar, Kr, Xe, Rn). Within this group the degree of inertness diminishes with increasing atomic number. All the other elements can form compounds with interatomic bonds of varying strength. These elements may be further divided according to their chemical properties into two groups, also unequal: *metals and metalloids*. All the elements of the A-subgroup, that is, all the elements to the left of the column of heavy univalent metals Cu, Ag, Au (Table I, Appendix 1), including the lanthanides and actinides, are metals. Among the elements of the B-subgroups a gradual change in their properties is observed from metallic to metalloid, if in this part of the Periodic Table one goes from left to right or vertically downwards. The most typical metalloids are the group of elements in the upper right hand corner, that is F, O, Cl. In view of the gradualness of the transition it is impossible to draw a sharp line between metals and metalloids but such a dividing line would pass roughly along the diagonal leading from boron to iodine. Near the line are elements possessing typical semiconducting properties (Si, Ge, Se, Te).

The distribution of the elements among these groups is thus:

inert	6.
metals	79.
metalloids . . .	8.
intermediate . .	7.

The intermediate elements include, besides the semi-conductors just mentioned, H, B and C as well. In the Periodic Table the most numerous group is the metals, comprising about 80% of all the elements. The general division of the elements between these groups depends on the peculiarities of the electron structures of their atoms.

The Periodic Law as the Fundamental Law of Nature

The periodic law was discovered by Mendeleev as a result of profound analysis of the chemical and physical properties of the elements. The basis of Mendeleev's ordering of the elements was their atomic weight. Despite the fact that several elements were missing from this sequence and that for some, incorrect atomic weights had been obtained, the fundamental law—periodicity in the variations of their properties—was established by Mendeleev, and he saw in it one of the fundamental laws of nature.

Mendeleev's deep conviction of the truth of his newly discovered law was shown in his use of it to correct atomic weights and predict as yet undiscovered elements. For some of these elements he indicated all their basic properties. These elements, with the properties predicted by Mendeleev, were subsequently discovered and he himself witnessed the brilliant confirmation of a number of his prophecies.

The periodic law made it necessary to transpose some of the elements and so destroy the sequence of increasing atomic weights. At the present time three such deviations are known (18 Ar – 19 K, 27 Co – 28 Ni, 52 Ne – 53 I). The atomic weight is only approximately proportional to the atomic number and these deviations are caused by the existence of isotopes. Despite these deviations Mendeleev placed the elements in an order corresponding to decreasing atomic weight and, as we now know, to increasing atomic number.

The periodic law is one of the fundamental laws of nature and describes the physical and chemical properties of matter throughout the whole Universe: on Earth, on other planets, in the atmospheres of the Sun and the stars, in the particles of cosmic dust, etc. The envoys of distant worlds—meteorites, falling onto the Earth—provide direct confirmation of the identity of the range of elements and properties of matter in different regions of the Universe. The spectra of nebulae and stars testify to the existence on them of well-known elements. Helium, first detected on the Sun, was later discovered on the Earth. Spectroscopy was used to confirm the formation in the Sun's atmosphere of a set of the simplest radicals, sufficiently stable at high temperatures (C_2, CH, CN, NH, OH, SiH, SiF, MgH). In the atmospheres of the planets the spectroscope uncovers the presence of such molecules as carbon dioxide (Venus), methane (Saturn, Uranus, Neptune), ammonia (Jupiter) etc. Astronomers believe that on the nearest planet to us, Mars, there are higher forms of organised matter—

that is, plants. Mendeleev's law has given one of the first foundations for the hypothesis of the possibility of life on distant worlds. Indeed the theme of one science-fiction novel is how the first cosmonauts use the Periodic Table as a key to decipher the Martian alphabet.

Many great scientific generalisations, such as, for example, the cosmogenic hypothesis of Kant and Laplace, have had only a temporary importance and represent in the history of science only one of the transient stages of its development. The fundamental law of nature—the periodic law, discovered by Mendeleev—is the foundation of Natural Philosophy and will exist for ever, like the very Universe itself.

6. The Electron Structure of Atoms of Different Groups of Elements

Elements of the B-subgroups

The simplest structure is that of the atoms of the lightest elements—hydrogen and helium, in which the first electron shell is being filled. This shell consists of a 1*s*-subgroup (Fig. 36*a*). In the atom of helium this subgroup is completely filled, thus marking the termination of the first period. In the succeeding elements, starting with lithium, the next electron group is being filled. It consists of two subgroups. The electron structure of atoms belonging to elements of the second period is shown in Fig. 36*b*. The filling of the *s*- and *p*-cells ends with neon.

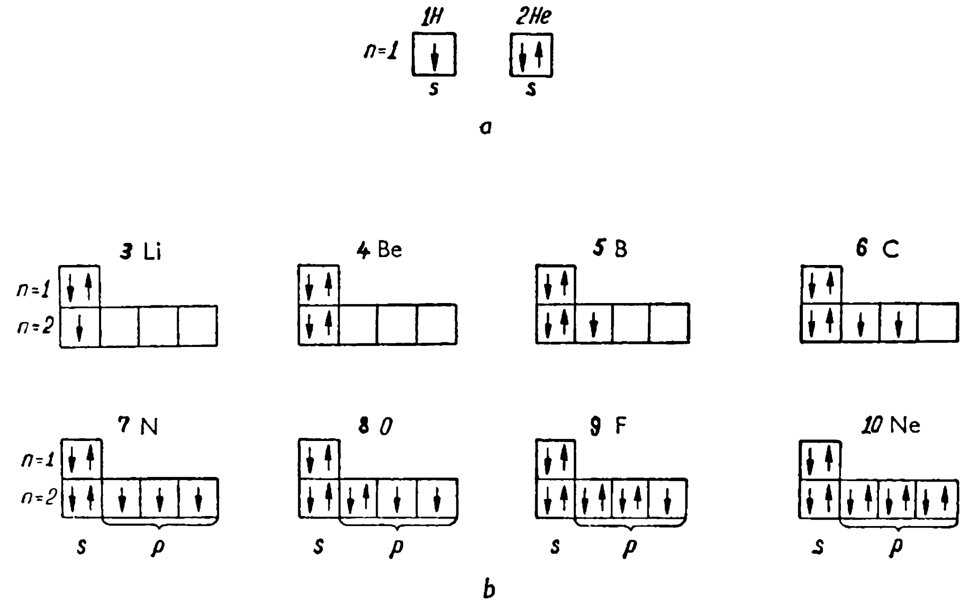


FIG. 36. Electron structure of atoms of (*a*) the first, (*b*) the second periods.

Table 9 gives the formulae for the electron configurations of the unfilled valence shell (upper line) and symbols of the atoms (lower line) for all the remaining elements which contain only *s*- and *p*-electrons. The last column gives, in the appropriate rows, the formulae for the completely filled inner groups. *s*- and *p*-valence electrons are found in atoms of the following elements: the alkali metals, alkaline earth metals and the elements of the B-subgroups. The electron configurations of the valence shell in these elements vary according to a simple law. In atoms of the alkali and alkaline earth elements the *ns*-cells are being filled. In elements of the B-subgroups first the *s*- and then the *p*-cells are being filled in succession, as shown in the schemes set forth above. The electron configurations of positive ions coincide with the electron configurations of atoms which are one or two groups further to the left (for instance, the configuration of the ion Mg^+ is the same as that of the atom Na, and similarly the ion Mg^{+2} corresponds to the atom Ne). The configurations of negative ions are the same as those of atoms one group to the right (the ion Cl^- has the configuration of the argon atom).

Table 9

Electron Configurations of Atoms with *s*- and *p*- Valence Electrons

	s	s^2	s^2p	s^2p^2	s^2p^3	s^2p^4	s^2p^5	s^2p^6	
n-2	Li^3	Be^4	B^5	C^6	N^7	O^8	F^9	Ne^{10}	$1s^2$
3	Na^{11}	Mg^{12}	Al^{13}	Si^{14}	P^{15}	S^{16}	Cl^{17}	A^{18}	$2s^22p^6$
4	K^{19}	Ca^{20}							$3s^23p^6$
4	Cu^{29}	Zn^{30}	Ga^{31}	Ge^{32}	As^{33}	Se^{34}	Br^{35}	Kr^{36}	$3d^{10}$
5	Rb^{37}	Sr^{38}							$4s^24p^6$
5	Ag^{47}	Cd^{48}	In^{49}	Sn^{50}	Sb^{51}	Te^{52}	I^{53}	Xe^{54}	$4d^{10}$
6	Cs^{55}	Ba^{58}							$5s^25p^6$
6	Au^{79}	Hg^{80}	Tl^{81}	Pb^{82}	Bi^{83}	Po^{84}	At^{85}	Rn^{86}	$4f^{14}5d^{10}$
7	Fr^{87}	Ra^{88}							$6s^26p^2$
	$^2S_{1/2}$	$^1S_0$	$^2P_{1/2}$	$^3P_0$	$^4S_{3/2}$	$^3P_2$	$^2P_{3/2}$	$^1S_0$	

The Transition Metals

In the more complicated cases arising in the succeeding periods (transition metals, rare earths, actinides), the theory cannot yet uniquely determine the distribution according to subgroups *d* and *s*, or *f*, *d* and *s*, of valence electrons possessing almost identical energies, so experimental data, mainly spectroscopic, are used to determine the electron configurations of the atoms. Fig. 37 shows the filling of quantum cells in elements of the fourth period. Table 10 gives the electron configurations and symbols of the atoms of elements in which a *d*-subgroup is being built up.

These elements are divided into sets according to their properties, and they form the large group of transition metals. Here the closeness of the d^ps^2-, $d^{p+1}s$- and d^{p+2}-states is apparent. The basic arrangement is the first of these and we find it in the majority of transition elements. The atom of Cr has the configuration $3d^54s$, when one of the *s*-electrons crosses over to the 3*d*-level. During this process all the *d*-cells in the Cr atom are simultaneously filled up and the atom acquires spherical symmetry. A similar transition occurs also in the Cu atom and in metals of the palladium series: Nb, Mo, Tc and Ru. Transition of both *s*-electrons to the *d*-level happens in Pd and Ir, whose atoms have only *d*-electrons in the ground state. The electron configurations of ions of the transition metals are not the same as the electron configurations of atoms in the neighbouring groups to the left. Thus the ion V^+ has the configuration $3d^4$, not $3d^24s^2$ as in the atom of Ti; the ion Fe^+ is $3d^64s$, not $3d^54s^2$ as in Mn. As Landau and Lifschitz have observed, all the ions found in the natural

	3*d*					4*s*	4*p*			
19 K						↓				
20 Ca						↓↑				
21 Sc	↓					↓↑				Transitional metals of the 4th period
22 Ti	↓	↓				↓↑				
23 V	↓	↓	↓			↓↑				
24 Cr	↓	↓	↓	↓	↓	↓				
25 Mn	↓	↓	↓	↓	↓	↓↑				
26 Fe	↓↑	↓	↓	↓	↓	↓↑				
27 Co	↓↑	↓↑	↓	↓	↓	↓↑				
28 Ni	↓↑	↓↑	↓↑	↓	↓	↓↑				
29 Cu	↓↑	↓↑	↓↑	↓↑	↓↓	↓				
30 Zn	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑				

FIG. 37. Electron structure of atoms of elements of the fourth period.

	4*f*							5*s*	5*p*			5*d*					6*s*	
55 Cs								↓↑	↓↑	↓↑	↓↑						↓	
56 Ba								↓↑	↓↑	↓↑	↓↑						↓↑	
57 La								↓↑	↓↑	↓↑	↓↑	↓					↓↑	
58 Ce	↓	↓						↓↑	↓↑	↓↑	↓↑						↓↑	
59 Pr	↓	↓	↓					↓↑	↓↑	↓↑	↓↑						↓↑	
60 Nd	↓	↓	↓	↓				↓↑	↓↑	↓↑	↓↑						↓↑	
61 Pm	↓	↓	↓	↓	↓			↓↑	↓↑	↓↑	↓↑						↓↑	
62 Sm	↓	↓	↓	↓	↓	↓		↓↑	↓↑	↓↑	↓↑						↓↑	
63 Eu	↓	↓	↓	↓	↓	↓	↓	↓↑	↓↑	↓↑	↓↑						↓↑	max. d (analogue of Mn)
64 Gd	↓	↓	↓	↓	↓	↓	↓	↓↑	↓↑	↓↑	↓↑	↓					↓↑	
65 Tb	↓↑	↓	↓	↓	↓	↓	↓	↓↑	↓↑	↓↑	↓↑	↓					↓↑	
66 Dy	↓↑	↓↑	↓↑	↓	↓	↓	↓	↓↑	↓↑	↓↑	↓↑						↓↑	
67 Ho	↓↑	↓↑	↓↑	↓↑	↓	↓	↓	↓↑	↓↑	↓↑	↓↑						↓↑	
68 Er	↓↑	↓↑	↓↑	↓↑	↓↑	↓	↓	↓↑	↓↑	↓↑	↓↑						↓↑	
69 Tm	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓	↓↑	↓↑	↓↑	↓↑						↓↑	
70 Yb	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑						↓↑	max. d
71 Lu	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓					↓↑	
72 Hf	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓	↓				↓↑	
73 Ta	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑	↓	↓	↓			↓↑	

FIG. 38. Electron structure of the lanthanides.

state in crystals and solutions contain in their valence shells only *d*- or *f*- (but not *s*- and *p*-) electrons. Thus iron occurs only in the form of ions Fe^{+2} and Fe^{+3} with configurations $3d^6$ and $3d^5$ respectively.

Lanthanides and Actinides

The electron configurations of the lanthanides are shown in Table 11 and in Fig. 38. In these elements, starting from cerium, the 4*f*-shell is being filled up, and a competition for electrons is taking place between the 4*f*- and 5*d*-states.

The least accurately known electron configurations are those of atoms of the actinides, in which there is competition for electrons between the 5*f*-, 6*d*- and 7*s*-states. Table 12 gives the electron configurations obtained by means of magnetic measurements and analysis of the chemical properties of compounds of the heavy elements. The most accurate figures are printed in bold type. According to the magnetic data, the overlapping of the 6*d*- and 5*f*-levels occurs when almost half the

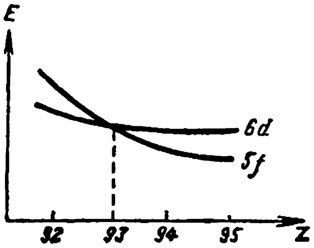


FIG. 39. Position of 5*f*- and 6*d*-levels.

Table 10

Electron Configurations of Atoms of Transition Metals

Nickel group

Core A +	Sc^{21}	Ti^{22}	V^{23}	Cr^{24}	Mn^{25}	Fe^{26}	Co^{27}	Ni^{28}
	$3d4s^2$ $^2D_{3/2}$	$3d^24s^2$ $^3F_2$	$3d^34s^2$ $^4F_{3/2}$	$3d^54s$ $^7S_3$	$3d^54s^2$ $^6S_{5/2}$	$3d^64s^2$ $^2D_4$	$3d^74s^2$ $^2F_{9/2}$	$3d^84s^2$ $^3F_4$

Palladium group

Core Kr +	Y^{39}	Zr^{40}	Nb^{41}	Mo^{42}	Tc^{43}	Ru^{44}	Rh^{45}	Pd^{46}
	$4d5s^2$ $^2D_{3/2}$	$4d^25s^2$ $^3F_2$	$4d^45s$ $^6D_{1/2}$	$4d^55s$ $^7S_3$	— —	$4d^75s$ $^5F_5$	$4d^85s$ $^4F_{9/2}$	$4d^{10}$ $^1S_0$

Platinum group

Core Xe +	La^{57}
	$5d6s^2$ $^2D_{3/2}$

Core Xe + $4f^{14}$ +	Lu^{71}	Hf^{72}	Ta^{73}	W^{74}	Re^{75}	Os^{76}	Ir^{77}	Pt^{78}
	$5d6s^2$ $^2D_{3/2}$	$5d^26s^2$ $^3F_2$	— —	$5d^46s^2$ $^5D_0$	$5d^56s^2$ $^6S_{5/2}$	— —	$5d^9$ $^2D_{5/2}$	$5d^96s$ $^3D_3$

Table 11

Electron Configurations of Neutral and Singly Ionised Atoms of Rare Earth Elements

Element	Symbol	Atomic number	Neutral atom						Ionized atom M^+: configuration and spectral term†
			Configuration†						
			$4f$	$5s$	$5p$	$5d$	$6s$	Spectral term	
Caesium	Cs	55	0	2	6	0	1	$^2S_{1/2}$	$^1S_0$
Barium	Ba	56	0	2	6	0	2	$^1S_0$	$^2S_{1/2}$
Lanthanum	La	57	0	2	6	1	2	$^2D_{3/2}$	$5d^2$ $^3F_2$
Cerium	Ce	58	(2	2	6	0	2	$^3H)$	$4f^26s$ $^4H_{7/2}$
Praesodymium	Pr	59	(3	2	6	0	2	$^4I)$	$4f^36s$ $^5I_4$
Neodymium	Nd	60	(4	2	6	0	2	$^3I)$	$4f^46s$ $^6I_{7/2}$
Promethium	Pm	61	((5	2	6	0	2	$^6H))$	$4f^66s$ $^8F_{1/2}$
Samarium	Sm	62	6	2	6	0	2	$^7F_0$	$4f^76s$ $^9S_4^0$
Europium	Eu	63	7	2	6	0	2	$^8S_{7/2}$	$4f^75d6s$ $^{10}D_{5/2}^0$
Gadolinium	Gd	64	7	2	6	1	2	$^9D_2$	
Terbium	Tb	65	((8	2	6	1	2	$^8H))$	
Dysprosium	Dy	66	((10	2	6	0	2	$^5I))$	
Holmium	Ho	67	((11	2	6	0	2	$^4I))$	
Erbium	Er	68	((12	2	6	0	2	$^3H))$	
Thulium	Tm	69	13	2	6	0	2	$^2F_{7/2}^0$	$4f^{13}6s$ $^3F_4^0$
Ytterbium	Yb	70	14	2	6	0	2	$^1S_0$	$4f^{14}6s$ $^2S_{1/2}$
Lutetium	Lu	71	14	2	6	1	2	$^2D_{3/2}$	$4f^{14}6s^2$ S_0
Hafnium	Hf	72	14	2	6	2	2	$^3F_2$	
Tantalum	Ta	73	14	2	6	3	2	$^4F_{3/2}$	

† Inner shells wholly filled and not shown in Table.

Table 12

Electron Configurations of Atoms of the Actinides

Element	Z	$5f$	$5f-6d$	$6s$	$6p$	$6d$	$7s$
Fr	87			2	6		1
Ra	88			2	6		2
Ac	89			2	6	1	2
Th	90			2	6	2	1
Pa	91			2	4	3	2
U	92			2	4	4	2
Np	93		5	2	4		2
Pu	94	5		2	4	1	2
Am	95	6		2	4	1	2
Cm	96	7		2	4	1	2
Bk	97	8		2	4	1	2
Cf	98	9		2	4	1	2

places in the 6*d*-subgroup have been filled, at which point the 5*f*-level becomes more favourable than the 6*d*-level (Fig. 39). The chemical data also point to a difference of properties between the pre-uranium and post-uranium elements. Th, Pa and U are more like the transition metals of the IVA, VA and VIA-subgroups than like the corresponding lanthanides. The question of the structure of the valence electron shell in atoms of heavy elements requires further study.

Problems

1. What kind of symmetry is found in the electron shell of a filled *p*-subgroup?
2. What form would a Periodic Table of 100 elements take if the electron levels were filled up according to the scheme of the single-electron atom? What would be the value of the magnetic moment of the atom with atomic number 100?

Chapter 3

THE QUANTUM THEORY OF INTERATOMIC AND INTERMOLECULAR INTERACTION

1. Interaction between Atoms and Electrons

Ionisation Potential of Atoms

The curve of first ionisation potential against atomic number (Fig. 34) shows the energy necessary to remove one of the outer electrons from an atom. The curve has an obviously periodic character. The smallest ionisation potentials are 3-5 eV and are for the alkali metals, where the atoms have only one electron in their valence shells. The largest ionisation potential (25·4 eV) is that for helium, which has a completely filled *K*-shell ($1s^2$-group). The same is true for the other noble gases, the atoms of which have 8-electron configurations, consisting of filled ns^2p^6-subgroups, in their valence shells. These configurations are characterised by great stability. The ionisation potential increases in proportion to the number of electrons in the valence shell of the atom, because of the increase in the nuclear charge, which raises the energy of interaction between the electrons and the nucleus. The mean distance of the electrons of a given subgroup from the nucleus is approximately the same, so the ionisation potential of helium is almost twice that of hydrogen; slightly less than twice because of the screening of the nucleus by the second *s*-electron. On the increasing sectors of the ionisation potential curve, small maxima and minima are observed; their appearance is connected with the finer details of the structure of the electron shell. Thus the ionisation potential maxima of the divalent elements Be, Mg, Zn, Cd, Hg are due to the fact that in these elements an s^2-subgroup is completed and the configuration formed by a pair of ns^2-electrons possesses relatively greater stability. The degree of stability of the valence pair of s^2-electrons grows with increasing atomic number, as is indicated by the increasing depth of the minima before the elements B, Al, Ga, In and Tl, where the *p*-subgroups start to be filled. The depth of the Be–B minimum is only 1·5 eV, whereas for Hg–Tl it is about 4 eV.† While the *p*-subgroups are less than half filled (B–N) the electron bond strength increases, but when the second half of the *p*-states begins to be filled the bond strength drops slightly. This is indicated by the minimum at oxygen and the stationary point in the section of the curve around sulphur. In the sector K–Zn two plateaux can easily be seen. They correspond to the filling of the first half and second half of the 3*d*-states. The filling of the first half of the 3*d*-states is complete with chromium, the element with atomic number 24.

---

† The great strength of a pair of s^2-electrons in heavy elements is the explanation for the decrease in valency (from 3 to 1) which is observed in a series of thallium compounds and in lead, and to a lesser extent tin, compounds (in both these latter cases from 4 to 2). (*Vide* Table II, Appendix 1.) This would seem also to explain the low heat of sublimation and low melting point of mercury.

Electron Affinity

When an electron approaches a neutral atom, forces of interaction which are mainly attractive arise between them. The energy of interaction varies with distance according to a different law from Coulomb's, being approximately as $1/r^4$. As shown by the solution of Schrödinger's equation, interaction obeying Coulomb's law would lead to the formation of an infinite number of allowed energy levels within the atom. But with an interaction which obeys the other law, the number of energy levels is finite and in particular cases they may even be absent. In the latter event negative ions are not formed. If there are in fact stable levels, then the atom tries to capture a *supernumerary* electron and turn itself into a negative ion. In these cases the energy with which the electron is captured, E, is called the *energy of affinity between the atom and the electron* or *electron affinity* and is positive. The greatest electron affinities are found in the halogen atoms (Table 13).

Table 13

Electron Affinities

Atom	E, eV	Atom	E, eV
F	4·1	I	3·3
Cl	3·7	O	3·0
Br	3·6		

When a halogen atom captures the additional electron a vacant place in the *p*-subgroup is filled up and a closed, stable 8-electron configuration is formed, like the configuration of the outer shell of an atom of a noble gas. The oxygen atom also possesses a high electron affinity but the addition of a second electron with the formation of a doubly charged negative oxygen ion O^{-2} is no longer advantageous and requires considerable energy. Negative electron affinity is predicted theoretically for atoms of the noble gases. Metallic atoms have vacant valence shells and many of them show a significant electron affinity (Li—0·5; Na—0·2 eV); this is especially great in the case of the heavy univalent metals (Cu ~ 1·0; Ag ~ 1·1; Au ~ 2·4 eV).

A high electron affinity is also found in certain polyatomic groups forming complex anions, for example NO_3^-, CO_3^{--}, SO_4^{--}, etc. At the same time some other polyatomic groups show affinity to a positive charge (proton), as for example in the well-known ammonium ion NH_4^+.

Electronegativity

In interactions between atoms their capacity to capture or give up a valence electron is very important. This capacity is measured by a quantity known as the electronegativity. The electronegativity is a relative property. Thus when atom B interacts with atom A it may capture an electron from it, but, conversely, when it interacts with atom C it may give up its electron. In this case the atoms are arranged according to electronegativity in the series A<B<C. The electronegativity can

be taken as approximately half the sum of the ionisation potential and the electron affinity energy: $x = \frac{1}{2}(I+E)$ (Table 14).

Table 14

Electronegativity of Atoms

H 2·1						
Li 1·0	Be 1·5	B 2·0	C 2·5	N 3·0	O 3·5	F 4·0
Na 0·9	Mg 1·2	Al 1·5	Si 1·8	P 2·1	S 2·5	Cl 3·0
K 0·8			Ge 1·8	As 2·0	Se 2·4	Br 2·8
Rb 0·8						I 2·5
Cs 0·8						

If the elements are arranged in order of electronegativity, then this series will begin with the most *electropositive* elements, that is, those most inclined to form positive ions. Such are the alkali metals, which have low ionisation potentials and insignificant electron affinities. At the other end of this series will be the most *electronegative* elements, that is, those most inclined to form negative ions. Such are the halogens, which have large ionisation potentials and the highest electron affinities. If the elements are arranged in order of electronegativity one can draw a boundary, albeit tentative, between those electropositive elements—the metals—and the electronegative elements—the metalloids. This boundary dividing the elements into two basic groups can be drawn according to their chemical properties, one of which is that in the majority of metallic chemical compounds such as salts positive ions (cations) are formed, whereas for metalloids the ions are negative (anions).

2. General Characteristics of the Forces of Interatomic Interaction

Electrical Nature of the Forces of Interaction

When atoms A and B approach one another, interaction begins between them. The interaction problem, even in the case of the simplest atoms, is extremely complicated, since it is necessary to consider the mutual behaviour of several particles (nuclei and electrons). To deal with it one must take account of the wave properties of microparticles, in the first instance of electrons, and solve the corresponding Schrödinger equation by approximate methods. In the general case it can be said that if the energy of the interacting atoms A and B is less than the total energy of the isolated atoms, then the forces of attraction prevail between them and they form a stable molecule AB. This interaction is described by the simplest chemical reaction equation $A+B = AB$.

If the energy of the system increases when the atoms approach one another, then the forces of repulsion predominate, and in this case the formation of a stable molecule is impossible. Let us consider what may be the nature of chemical interaction energy. The forces operating between atoms could be gravitational, magnetic or electric.

The potential energy of gravitational interaction is insignificantly small. For the heaviest atoms, of atomic weight 250,

$$U_{\mathrm{grav}} = -G\frac{(m_{\mathrm{H}}A)^2}{r} = -6{\cdot}7\times 10^{-8}\frac{(1{\cdot}66\times 250)^2\times 10^{-48}}{3\times 10^{-8}}\ \mathrm{erg} =$$
$$= 3{\cdot}9\times 10^{-44}\ \mathrm{erg} = -2{\cdot}4\times 10^{-32}\ \mathrm{eV}, \qquad \ldots [\mathit{3.1}]$$

where G is the gravitational constant and m_{H} is the mass of the hydrogen atom. Gravitational forces cannot possibly account for the formation of molecules and crystals.

In order to evaluate the magnetic energy, we shall find the interaction energy of two elementary magnetic dipoles, each equal to the Bohr magneton, and a distance 3 Å apart:

$$U_{\mathrm{mag}} = -\frac{2M_{\mathrm{B}}^2}{r^3} = -2\frac{(0{\cdot}928\times 10^{-20})^2}{27\times 10^{-24}}\ \mathrm{erg} = -7\times 10^{-6}\ \mathrm{eV} = -0{\cdot}2\ \mathrm{cal./mole} \qquad \ldots [\mathit{3.2}]$$

This energy is considerably greater than the gravitational, and in some cases can be measured. Thus in the ferromagnetic crystal of magnetite (Fe_3O_4) the magnetic energy, found from experimental data, ~0·1 cal./mole. However even the magnetic energy is considerably less than the mean binding energy of atoms in molecules and crystals.

The electrostatic interaction energy of two elementary electric charges a distance 3 Å apart is

$$U_{\mathrm{e.s.}} = -\frac{e^2}{r} = -\frac{(4{\cdot}8\times 10^{-10})^2}{3\times 10^{-8}} = 8\times 10^{-12}\ \mathrm{erg} = -5\ \mathrm{eV} \approx -220\ \mathrm{kcal./mole} \qquad \ldots [\mathit{3.3}]$$

This corresponds in order of magnitude to the energy of a chemical bond, representing the energy of electrical interaction between atoms. The chemical binding energy is comparable in magnitude to the mean binding energy of the outer electrons in atoms. The chemical bond is created by these outer electrons, which are called valence electrons. The measure of the capacity of an atom to take part in a chemical interaction is its *valency*.

The Three Basic Types of Chemical Bond

As was pointed out above, the chemically active elements are divided into two fundamental groups, the metals and the metalloids, the atoms of which have electropositive (+) and electronegative (−) properties respectively. Chemical interaction occurs not only between atoms belonging to the same group, but also between atoms belonging to different groups. There are three fundamental types of chemical bond, corresponding to the three different types of combinations of atoms, depending on their electronegativities: (− −), (+ +), (+ −). These bonds are the covalent, the metallic and the ionic (Table 15).

Table 15

Basic Types of Chemical Bond

Chemical nature of interacting atoms	Polarity of bond	Type of chemical bond	Symbols
Metalloid-metalloid	Homopolar	Covalent	− −
Metal-metal	Homopolar	Metallic	+ +
Metal-metalloid	Heteropolar	Ionic	+ −

The electrical nature of the interaction is the same in all these cases but, the character of the interaction varies substantially because of the difference in electron structure between atoms of different groups. The difference in the nature of the chemical interaction is shown by the considerable variation in the physical and chemical properties of simple and compound substances.

A peculiarity of the electron structure of metallic atoms is the partial filling of the outer (valence) shell. Thus in the atoms of the alkali metals, which begin new periods in the Periodic Table, there is one s-electron in the valence shell. Neighbouring p-levels remain free (Fig. 40a). In the atoms of alkaline earth metals, and also in the atoms of transition metals, including rare earth and heavy metals at the end of the Periodic Table, the valence shell has its s, p, d and f-levels respectively only partially filled; a large part of these sublevels remains free.

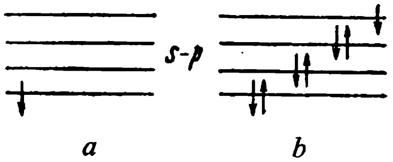


FIG. 40. Filling of the valence shells in atoms of alkali metals and halogens.

A peculiarity of the electron structure of the atoms of metalloids is that the valence shell is to a large extent filled up, and only a few levels in it remain unoccupied. Thus in halogen atoms only one level is vacant in the $s-p$-shell (Fig. 40b). In the elements of the oxygen subgroup two free levels remain, and in the nitrogen subgroup three. In elements of the carbon subgroup the $s-p$-shell is half filled, each atom having 4 electrons and 4 free levels in this shell. The next unoccupied energy levels in metalloid atoms lie too high by comparison with the chemical bond energy so any hope of using them for electron sharing when atoms approach one another is ruled out by energy considerations.

Intermediate Types of Chemical Bond

The basic types of chemical bond shown in Table 15 describe the limiting cases of chemical interaction. In reality, a large number of intermediate types of bond are possible, depending on the nature and electronegativity differences of the reacting atoms. In molecules and crystals formed by atoms of the same kind (homopolar bonds), the valence electrons are shared and move in orbits which take in several nuclei. This leads in the majority of cases to a symmetrical valence electron density distribution between neighbouring nuclei. In molecules and crystals formed by atoms with different electronegativities (heteropolar bond), the centre of gravity of the electron cloud is displaced to the side of the atom with the greater electronegativity. If there is a considerable difference in the electronegativities, the valence electron of the electropositive atom passes over almost completely to the shell of the electronegative atom and in this case a compound is formed with the maximum amount of ionic binding.

Structural Characteristics of Bonds and the Distribution of Electron Density

An atom may have different valence states in different chemical bonds. These differences include: (1) the number of valence electrons taking part in bonds (the different valencies of atoms of oxygen, sulphur, iron, lead etc.); (2) the number and multiplicity of the bonds formed by given atoms with neighbouring atoms; (3) the configurations of the bonds (for example, the pyramidal and plane bonds formed by the nitrogen atom); (4) the redistribution of the valence electron density and degree of polarity of the bonds. The properties of compounds depend on the previously mentioned properties of chemical bonds, so a knowledge of the character of chemical interaction is very important. The study of the properties of chemical compounds, especially their electrical, magnetic and optical properties, provides important material concerning the characteristics of the bond and enables us to understand the nature of chemical interaction. Structural research is especially valuable as it enables us to establish the structural characteristics of bonds: configurations, valence angles and interatomic distances. The configurations determine the coordination number k, that is, the number of nearest neighbour atoms, and the bond distribution symmetry given by the valence angles. The valence angles must be known if we are to determine the type of orbit hybridisation occurring among the valence electrons. The interatomic distances are important for determining the multiplicities of the bonds and their degree of polarity. The accurate determination of interatomic distances is possible by structural methods and also by spectroscopy, radiospectroscopy, and nuclear magnetic resonance. The latter methods are beginning to be widely used for research into the structure of molecules and crystals.

Apart from structural characteristics, the methods of X-ray and electron diffraction enable us to determine the fine details of the redistribution of electron density caused by differences in electronegativity between the atoms. This redistribution can lead to the excitation of new valence states and create a bond which is polar to a greater or lesser degree. Data on the polarity of the bond can be obtained by research into paramagnetic electron resonance and, in some cases, by X-ray spectroscopy.

3. The Quantum Theory of Interatomic Interaction

Overlapping of Atomic Wave Functions

Let us consider atoms a and b, the nuclei of which are at points A and B respectively (Fig. 41) on the straight line CC. Let ψ_a^{nlm} and $\psi_b^{n'l'm'}$ be the wave functions which describe the behaviour of the valence electrons in these atoms. The probability of finding an electron at a distance r from the nucleus, in a thin spherical shell of radii r and $r+dr$, is given, according to [*2.47*], by the spherical density $\sigma(r) = R^2(r)r^2dr$, where $R(r)$ is the radial part of the wave function. Fig. 41 shows the form of the radial electron density distribution functions for $2s$ (atom a) and $2p$ (atom b) electrons.

The position of the principal maximum of each curve $\sigma(r)$ gives the mean distance at which the electron is most likely to be found in the atom. This distance (corresponding to the radius of the Bohr orbit in the old treatment) generally gives the energy. For atoms a and b, the principal maxima are at distances r_a and r_b from the respective nuclei. The quantities r_a and r_b may be called the atomic radii of the *free* atoms. These atomic radii do not vary much over all atoms in the Periodic Table, ranging from 0·5 Å for the lightest atom, hydrogen, to ~1·5 Å for multi-electron atoms. With increasing r, the curves $\sigma(r)$ rapidly fall away exponentially and approach the x-axis asymptotically.

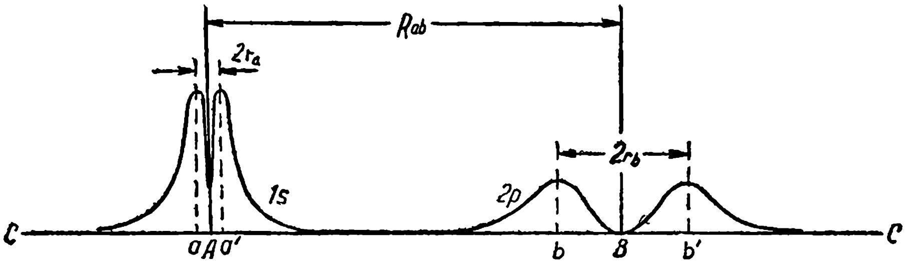


FIG. 41. Radial electron density distribution for atoms a and b.

We shall give the name interatomic distance to the distance R_{ab} between the nuclei of the atoms and so long as this distance is several times greater than the sum of the atomic radii the atoms, being electrically neutral systems, do not interact with one another. When the atoms approach to within a distance of the order of the sum of the atomic radii, *overlapping* of their wave functions begins: first, the wave functions of the outer electrons overlap, then when the principal maxima overlap, that is when the atoms approach to within distances

$$R_{ab} \sim r_a' + r_b, \qquad \ldots [3.4]$$

strong perturbation of the motion of these electrons (called *valence* electrons) begins under the influence of the fields of the neighbouring atoms.

When the valence electrons of both atoms are in the space between the nuclei, the electric fields acting on these electrons from both atoms are roughly the same and the valence electrons of the two atoms can be shared. This means that the characteristic movements of the valence electrons are altered as compared with their movements in the free atoms. The valence electrons begin to move in the field of both atoms and their motion will be described by a new shared wave function ψ, which we shall call the *molecular* wave function. Sharing of valence electrons is accompanied by redistribution of the electron density and a change in the energy $\mathscr{E}$ of the system as compared with the total energy of the two atoms, $\mathscr{E}_a + \mathscr{E}_b$. The quantum character is preserved in a system of two or more atoms. As in atoms, this shows in the discreteness of the energy levels of the system.

If the overlapping of the wave functions is accompanied by a reduction in the energy of the system, $\mathscr{E} < \mathscr{E}_a + \mathscr{E}_b$, then forces of attraction must arise between the atoms. When the atoms approach closer, these forces will give way to the rapidly increasing forces of nuclear repulsion. When the nuclei are separated by a certain distance, corresponding to the minimum system energy $\mathscr{E}_{\min} = E$, the forces of attraction will balance the forces of repulsion. This indicates that the system is in a state of *equilibrium* and means that a molecule has come into being which is qualitatively different from the initial set of two atoms.

> A molecule is not the additive sum of its constituent atoms. The creation of qualitatively new substances is well known from a wealth of experimental chemical data. Thus for example, when two very active chemical substances interact—such as the alkali metal sodium and the poisonous gas chlorine—a new and chemically inert substance appears—sodium chloride (common salt), with properties completely unlike those of the initial components.

The forces of attraction which arise when atoms approach one another and their wave functions overlap are often called the forces of cohesion, and are forces of chemical affinity, or *chemical binding* forces. The radius of action of these forces is, to a first approximation, the sum of the radii of the free atoms and lies between the limits ~1 and 3Å. The chemical binding energy (the cohesive energy) is equal to the difference between the total energy of the free atoms and the energy of the molecule, $(\mathscr{E}_a+\mathscr{E}_b)-\mathscr{E}_{\min}$. If this difference is positive, then the formation of the molecule AB from the free atoms is accompanied by the evolution of energy. Under these conditions a chemical reaction $A+B = AB$[†] can take place.

Schrodinger's Equation for a System of Atoms

Molecules and crystals are sets of nuclei and electrons. We shall denote nuclei by large, and electrons by small, italic letters. Fig. 42 shows the arrangement of nuclei and electrons at a certain instant of time. The radii vectores R_{JK} joining the nuclei to one another are drawn as thick lines, and the set of these radii vectores forms the framework of the molecule or crystal. The radii vectores r_{iK} joining the electrons to the nuclei are drawn as thin lines, and the radii vectores r_{ik} joining the electrons to one another, as dotted lines. We shall consider the stationary states of the system described by Schrödinger's time-independent equation

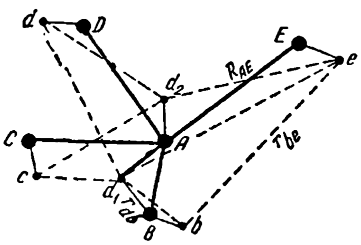


Fig. 42. System of nuclei and electrons.

$$H\psi = \mathscr{E}\psi, \qquad \ldots [3.5]$$

[†] Whether or not stable molecules of a compound AB are formed under the given conditions depends on the binding energy of the components in molecules (or crystals). If the formation of molecules and crystals from the pure components is accompanied by a considerable increase in energy, then the more stable state corresponds to a heterogeneous mixture A + B of the components (or to a homogeneous mixture of the corresponding gases). In some cases, even when there is a predisposition to form molecules of a compound rather than molecules of the individual components, the system can nevertheless be a mixture of gases and remain for a long time in the metastable state, if this is the case it is necessary to expend an amount of energy greater than the mean heat energy of the molecules to start the reaction (for example, an unstable mixture of hydrogen and oxygen).

where H is the Hamiltonian operator of the system

$$H = K+U. \qquad \ldots [3.6]$$

Here K is the kinetic energy operator of the system, and U is its potential energy. The kinetic energy operator for the system under consideration can be written in the form

$$-K = \frac{h^2}{8\pi^2}\sum_J \frac{1}{M_J}\nabla_J^2 + \frac{h^2}{8\pi^2 m}\sum_i \nabla_i^2. \qquad \ldots [3.7]$$

Here M_J is the mass of the J-th nucleus, and m is the mass of an electron.

The potential energy of the system is equal to its total electrostatic energy

$$U = -\frac{e^2}{2}\sum_i\sum_K \frac{Z_K}{r_{iK}} + \frac{e^2}{2}\sum_i\sum_k \frac{1}{r_{ik}} + \frac{e^2}{2}\sum_J\sum_K \frac{Z_J Z_K}{R_{JK}}, \qquad \ldots [3.8]$$

where Z_J and Z_K are the charges on the corresponding nuclei. The factor $\frac{1}{2}$ in [8] appears because each radius vector is counted twice in the double summation. The first term in [8] is the energy of attraction of the electrons towards the nuclei; the second and third terms are the energies of repulsion of the electrons and nuclei respectively. Introducing [7] and [8] into [5], we obtain in its general form Schrödinger's equation for the system

$$\sum_J \frac{1}{M_J}\nabla_J^2\psi + \frac{1}{m}\sum_i \nabla_i^2\psi + \frac{8\pi^2}{h^2}\left(\mathscr{E} + \frac{e^2}{2}\sum_i\sum_K \frac{Z_k}{r_{iK}} - \frac{e^2}{2}\sum_i\sum_k \frac{1}{r_{ik}} - \frac{e^2}{2}\sum_J\sum_K \frac{Z_J Z_k}{R_{JK}}\right)\psi = 0, \qquad \ldots [3.9]$$

where ψ is the total wave function of the system, depending on the coordinates of all the particles in the system

$$\psi = \psi(x_1, y_1, z_1;\ x_2, y_2, z_2; \ldots\ x_i, y_i, z_i, \ldots;\ X_1, Y_1, Z_1, \ldots;\ X_J\ Y_J Z_J; \ldots),$$

where x_i, y_i, z_i are the coordinates of the i-th electron, and X_J, Y_J, Z_J those of the J-th nucleus. If we impose on the wave function ψ of the molecule, conditions suggested by its physical meaning (continuity, boundedness, single-valuedness), then equation [9], like Schrödinger's equation for the hydrogen atom, has solutions only for strictly defined discrete values of the parameter $\mathscr{E}$ of the equation. These values determine the energy levels of the molecule (or crystal). The square of the modulus of the wave function gives the probability of finding the particles in different regions of the system.

The great difference between the masses of electrons and nuclei and the corresponding difference between their degrees of inertia mean that the motions of electrons and nuclei are fundamentally different in character. In molecules and crystals, the electrons trace out curved paths just as in atoms, whereas the nuclei, which in atoms may be thought of as at rest, perform oscillatory motions about their positions of equilibrium. In an interval of time during which the electrons complete a large number of revolutions about the nuclei, the latter are displaced only a very slight distance. Thus, the forces acting on the nuclei may be averaged and considered as constant. When the system is in equilibrium these forces vanish; this corresponds to the equilibrium position of the nuclei. During displacement from the equilibrium positions, restoring forces begin to act on the nuclei, tending to return the latter to the equilibrium positions, and for small displacements, these forces are proportional to the displacements. Under their influence the nuclei perform harmonic oscillations with amplitudes which are small by comparison with the internuclear distances, and with frequencies ν of the order of 10^{13} sec.$^{-1}$ For the distances between nuclei in molecules and crystals, we shall take the distances between their equilibrium positions. The oscillatory motions of the nuclei are quantised. Owing to the zero point energy of the quantum oscillator, the nuclei, even at absolute zero, never come to rest in their equilibrium positions.

The configuration of atomic nuclei in the molecule or crystal does not vary with the passage of time:

$$R_{JK} = \text{const.} \ldots \qquad \ldots [3.10]$$

This enables us to speak of the definite *atomic structure of a molecule or crystal.* This structure can be indicated most simply by drawing the vectors linking adjacent nuclei (atoms). These vectors correspond to the lines of action of the chemical bonds, and show which atoms are directly linked with one another. The assemblage of lines of action of the bonds forms, as a rule, a rigid framework for the molecule (or crystal) and the electrons move within the limits of this framework. The electrons belonging to individual atoms move around their own nuclei, while the shared valence electrons take in several nuclei in their motions. As a result of the steady motion of the electrons in the molecule (or crystal) a definite electron density distribution is set up, and this plays an essential rôle in determining the properties of the molecule (or crystal). Thus for example, in heteropolar compounds, the redistribution of electron density gives rise to electric dipole moments, and they account for the special electrical properties of the corresponding substances. The configuration of atomic nuclei and the distribution of electron density in a molecule (or crystal) determine its structure (sometimes called its chemical structure).

The atomic arrangement of a substance can be established experimentally by means of structural methods of investigation (X-ray-, electron- or neutron diffraction) and by the study of the magnetic, electrical, optical and other properties of the corresponding molecules and crystals. Therefore in considering questions of interatomic interaction we shall rely on experimental data and regard the atomic structure as known. The expediency of this becomes more clear when one reflects that in a number of cases isomerism, allotropy and polymorphism occur; in other words, a single chemical substance may exist in various different structural forms.

The above argument shows that equation [*9*] can be simplified. Thus relation [*10*] leads to the result that the differential part of equation [*9*] which relates to the nuclei, and is expressed by the first term, vanishes. The last term in brackets, which expresses the energy of electrostatic repulsion of the nuclei, is a constant:

$$\frac{e^2}{2}\sum_J\sum_K \frac{Z_J Z_k}{R_{JK}} = C = \text{const.} \qquad \ldots [3.11]$$

Thus in Schrödinger's equation for molecules and crystals we need keep only the terms relating to the electron part:

$$\sum_i \nabla_i^2\psi + \frac{8\pi^2 m}{h^2}\left(\mathcal{E} + \frac{e^2}{2}\sum_i\sum_K \frac{Z_K}{r_{iK}} - \frac{e^2}{2}\sum_i\sum_k \frac{1}{r_{ik}}\right)\psi = 0. \qquad \ldots [3.12]$$

The molecular wave function will also be correspondingly simplified; it will now be a function only of the coordinates of the electrons:

$$\psi = \psi(x_1 y_1 z_1; \ldots x_i, y_i, z_i, \ldots). \qquad \ldots [3.13]$$

We should note that the rejected constant C—the energy of nuclear repulsion—in the expression for the potential energy will have to be reincluded later when the total energy $\mathscr{E}$ is calculated, since the total potential energies of the molecules and crystals are normalised in such a way that $U \to 0$ when the particles are infinitely far from one another.

Despite the simplifications, equation [*12*] is so complicated that mathematical difficulties make it insoluble. For the simplest case, that of the molecular ion of hydrogen H_2^+ (Fig. 43), which contains 3 particles—2 protons and 1 electron—equation [*12*] takes the form

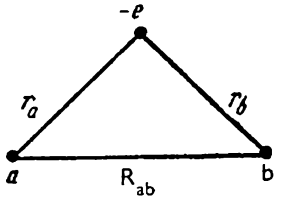


FIG. 43
Molecular ion of hydrogen.

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}\left(\mathscr{E} + \frac{e^2}{r_a} + \frac{e^2}{r_b}\right)\psi = 0. \qquad \ldots [3.14]$$

However even for this equation which relates to the simplest molecule, the exact solution is, for the reason stated, unknown. Thus approximate methods become of great value in the solution of the problem of interatomic interaction.

The Variational Method

If we multiply both sides of equation [*5*] by ψ^* and integrate throughout the whole volume of the system we obtain for its energy the expression

$$\mathscr{E} = \frac{\int \psi^* H \psi \, d\tau}{\int |\psi|^2 d\tau}. \qquad \ldots [3.15]$$

In this equation H, the Hamiltonian of the system, is known, and in order to calculate the energy we have to know the wave function ψ. In so far as the exact form of this function cannot be found by direct solution of Schrödinger's equation, we shall try to pick out approximate values for the molecular wave function, starting from the general physical conditions of the problems.

We may consider the best approximate molecular function ψ out of the given class of functions to be that which makes the total energy of the system, found by formula [*15*], a minimum. We can find the minimum energy of the system by varying $\mathscr{E}$ with respect to a parameter which characterises the chosen class of wave functions. The variation of the energy of the system can be continued further by introducing new and physically meaningful parameters which improve the chosen approximate expression for the wave function. The method of successive approximations may reasonably be used until the minimum energy of the system found by it has reached a practically stable value. Variational methods were successfully used long ago to solve a series of problems in classical mechanics and optics.

To apply the variational method in the present case one must find an approximate expression for the molecular wave function. In a sufficiently general form this can be achieved by means of the method of atomic orbitals (Heitler–London) or the method of molecular orbitals (Hund–Millikan–Herzberg).

The Method of Atomic Orbitals

Let us denote the atoms forming the molecule by the letters

$$a, b, \dots, i, \dots n \qquad \dots [3.16]$$

The valence electrons of the atoms will be denoted by the figures

$$1, 2, \dots, k, \dots n$$

For simplicity we shall consider the case of monovalent atoms, the number of which is equal to the number of valence electrons. We first extend the molecule until the atoms forming it are separated from one another by distances so large that interaction between them can be ignored. The behaviour of the valence electrons in atoms in this state, which we shall take as the initial state of the system, is described by the unperturbed atomic wave functions

$$\psi_a(1), \psi_b(2), \dots, \psi_k(k), \dots \psi_n(n). \qquad \dots [3.17]$$

A valence electron of each atom is linked with a given atom. The wave function ψ_1 of the assemblage of non-interacting atoms is equal to the product of the atomic wave functions

$$\psi_1 = \psi_a(1)\,\psi_b(2)\dots\psi_k(k)\dots\psi_n(n) = \prod_{k=1}^{n} \psi_k(k). \qquad \dots [3.18]$$

In the light of the probability interpretation of the wave function, formula [*18*] may be considered as a consequence of the theorem that the probability of a compound event is equal to the product of the probabilities of the component events, if the latter are independent of one another. The truth of this result may easily be verified by substituting the expression into Schrödinger's equation for the assemblage of atoms, whose Hamiltonian is equal to the sum of the Hamiltonians of the individual atoms, $H = H_1 + H_2 + \dots$, and whose total energy is equal to the sum of the total energies of the atoms, $\mathscr{E} = \mathscr{E}_1 + \mathscr{E}_2 + \dots$

When the atoms approach one another and a molecule is formed, the valence electrons become shared and move in the field of several atoms. An electron originally belonging to the first atom will after a certain time be near the second atom, and then the third and so on. The first atom will acquire a new valence electron in place of the one which has left it and gone to another atom. As a result there will be a redistribution of valence electrons, which will be described by a wave function ψ_2, analogous to [*18*], but with the electrons rearranged

$$\psi_2 = \psi_a(2)\cdot\psi_b(1)\ldots\psi_k(l)\ldots\psi_n(k) = \prod_{k=1}^{n}\psi_k(l). \qquad \ldots[3.19]$$

Corresponding to the new distributions of valence electrons between the atoms, we obtain a series of molecular wave functions differing from one another in the permutations of the valence electrons. Thus, the molecule is described by a set of molecular wave functions

$$\psi_1, \psi_2, \psi_3, \ldots, \psi_m, \ldots, \qquad \ldots[3.20]$$

each of which is an approximate particular solution of Schrödinger's equation for the molecule. The approximate character is connected with the fact that these solutions take into account only the sharing of the valence electrons, the 'electron exchange', and not the perturbations of the atomic wave functions ψ_k under the mutual influence of the atoms, since [*18*] includes only the unperturbed atomic wave functions.

The general solution of Schrödinger's equation can be written in the form

$$\psi = c_1\psi_1+c_2\psi_2+\ldots+c_m\psi_m+\ldots+c_n\psi_n, \qquad \ldots[3.21]$$

where the constant coefficients $c_1, c_2, \ldots, c_m \ldots$ define the degree of participation of each of the functions [*20*] in the general solution. The values of these coefficients may be found by the variational method.

The Method of Molecular Orbitals

There is a molecular wave function ψ, often called the 'molecular orbital', which describes the motions of an electron in the general field of two atoms a and b. In the region of atom a the electron experiences the predominant influence of that atom's nucleus and of the electrons surrounding its nucleus. So the molecular function ψ must be similar to the atomic function ψ_a. In the region of atom b the electron experiences the predominating influence of atom b's nucleus and surrounding electrons. Consequently the function ψ must resemble the atomic function ψ_b. Both demands can be satisfied by assuming as a first approximation that the molecular wave function ψ is a superposition (sum) of unperturbed atomic functions taken each with its appropriate weight

$$\psi = c_a\psi_a+c_b\psi_b = N(\psi_a+\lambda\psi_b), \qquad \ldots[3.22]$$

where c_a and c_b are constant coefficients measuring the degree of participation of each of the atomic orbitals. These coefficients may be expressed in terms of constants N and λ. Wave functions of type [*22*] describe the motion not only of valence electrons, but also of the electrons within atomic structures. In the latter case we must take, for the structure of atom a, $c_b = 0$, and for that of atom b, $c_a = 0$. Then the molecular functions transform into the corresponding atomic functions. The coefficients c_a and c_b in the general case are connected by the relationship

$$c_a^2+c_b^2+2c_ac_bS = 1 \qquad \ldots[3.23]$$

and there is a corresponding relationship between N and λ,

$$N = (1+\lambda^2+2\lambda S)^{-\frac{1}{2}}. \qquad \ldots [3.24]$$

These are obtained from the normalisation condition for the molecular wave function, $\int |\psi|^2 d\tau = 1$. Formulae [*23*] and [*24*] contain the quantity S,

$$S = \int \psi_a \psi_b d\tau, \qquad \ldots [3.25]$$

which is called the overlap integral.

Substituting [*22*] into [*15*], we obtain

$$\mathscr{E} = \frac{c_a^2 E_a + 2c_a c_b \beta + c_b^2 E_b}{c_a^2 + 2c_a c_b S + c_b^2}. \qquad \ldots [3.26]$$

In formula [*26*]

$$\beta = \int \psi_a H \psi_b d\tau \qquad \ldots [3.27]$$

is the so-called resonance integral. The quantities

$$E_a = \int \psi_a H \psi_a d\tau, \quad E_b = \int \psi_b H \psi_b d\tau \qquad \ldots [3.28]$$

are the portions of molecular energy connected with the atomic functions (but not the atomic levels, in as much as [*28*] contains the Hamiltonian of the molecule, not those of the individual atoms). If the atoms are sufficiently far apart E_a and E_b tend to the values of the energies of the unperturbed atomic levels E_a^0, E_b^0.

Calculation of Atomic Interaction by the Method of Molecular Orbitals

We must choose the values of the coefficients c_a and c_b in [*22*] to be such that the energy $\mathscr{E}$ of the molecule, expressed by formula [*26*], shall be a minimum. The extremum conditions for the energy take the form

$$\frac{\partial \mathscr{E}}{\partial c_a} = 0 \quad \text{and} \quad \frac{\partial \mathscr{E}}{\partial c_b} = 0.$$

These conditions lead to the two equations

$$\left.\begin{aligned} c_a(E_a-E)+c_b(\beta-ES) &= 0 \\ c_a(\beta-ES)+c_b(E_b-E) &= 0 \end{aligned}\right\}, \qquad \ldots [3.29]$$

in which E denotes the extremal value of $\mathscr{E}$. Non-zero solutions of equation [*29*] are possible if the determinant formed from the coefficients of these equations vanishes

$$\begin{vmatrix} E_a-E, & \beta-ES \\ \beta-ES, & E_b-E \end{vmatrix} = 0.$$

In expanded form this gives the quadratic equation

$$(E-E_a)(E-E_b)-(\beta-ES)^2 = 0, \qquad \ldots [3.30]$$

from which E can be determined. Knowing E, we can obtain the ratio $\lambda = c_b/c_a$ of the coefficients from either of equations [*29*].

For symmetrical diatomic molecules of the type A_2, formed from identical atoms, the above relationships simplify. If $E_a = E_b$, then from [*29*] we have

$$\lambda^2 = 1 \quad \text{and} \quad \lambda = \pm 1. \qquad \ldots [3.31]$$

This means that the moduli $|c_a|$ and $|c_b|$ of the coefficients in [*22*] are the same. This is a consequence of the symmetry of the molecule.

From [*31*] it follows that when two atomic functions are superposed two molecular functions are formed, the symmetric

$$\psi_1 = \psi_a + \psi_b \qquad \ldots [3.32]$$

and the antisymmetric

$$\psi_2 = \psi_a - \psi_b. \qquad \ldots [3.33]$$

From [*30*] we find the energies of the diatomic molecule in these two states

$$E_1 = \frac{E_a + \beta}{1+S} \quad \text{and} \quad E_2 = \frac{E_a - \beta}{1-S}. \qquad \ldots [3.34]$$

Formulae [*34*] show that when the atoms approach one another and the atomic wave functions overlap ($\beta \neq 0$, $S \neq 0$), the atomic levels split up. From the doubly degenerate atomic levels (that is, coinciding in magnitude, $E_a = E_b$) two molecular levels are formed, the energy of one of them being diminished, while that of the other is increased as compared with the energy of each unperturbed atomic level. The decrease in one of the molecular levels shows that when the atoms approach one another the total energy of the system may be diminished; this corresponds to the appearance of forces of attraction between the atoms. Since the resonance integral β is negative, the lower molecular level is described by the symmetric combination of the atomic orbitals $(\psi_a + \psi_b)$, and the upper by the antisymmetric $(\psi_a - \psi_b)$. The overlap integral $S < 1$, so the upper molecular level rises somewhat more rapidly than the lower one falls: $E_a - E_1 < E_2 - E_a$. The lowering of the lower molecular level as the atoms approach one another continues at first because S and β are increasing, but later the molecular energy begins to increase because of the rapid growth of the nuclear repulsion energy as the atoms come even closer together.

Let us consider the graphical solution of equation [*30*] for an asymmetrical molecule AB. The L.H.S. of [*30*], $f(E)$ say, represents the equation of a parabola, the graph of which is shown in Fig. 44 for the case $E_a < E_b$. For energy values lying in the interval $E_a \leqq E \leqq E_b$, $f(E)$ takes negative values. For values of E with greater modulus, $f(E) > 0$. The roots of equation [*30*], E_1 and E_2, obtained from the intersections of the parabola with the x-axis, lie outside the energy interval E_a to E_b. This is shown also by the analytic solutions of [*30*]. If we suppose that one of the roots is near E_a and the other near E_b, and substitute first E_a and then E_b in the first term of [*30*] except in the factor which would otherwise vanish, we obtain approximate solutions in the form

$$E_1 = E_a - \frac{(\beta - ES)^2}{E_b - E_a}; \quad E_2 = E_b + \frac{(\beta - ES)^2}{E_b - E_a}. \qquad \dots [3.35]$$

The root E_1 lies to the left, and E_2 to the right, of the interval $E_b - E_a$. This result does not depend on the values of E_a, E_b, β and S, provided β and S do not vanish simultaneously.

As was shown above, the overlapping of atomic functions ψ_a and ψ_b leads to the formation of two molecular wave functions which describe the state of a molecule with different energies. The energy of one state is less, and of the other greater, than the energies of the original atomic states. The molecular levels are filled up with electrons in exactly the same way as atomic levels. Each molecular state accommodates at most two electrons with opposite spins. This consequence of the Pauli principle explains the empirical rule of Lewis (1916), who established the rôle of electron pairs in the formation of covalent bonds: the lower molecular levels are filled first, and then those of higher energy. The diagram of molecular levels shows why some atoms form stable molecules, while others neither form stable molecules nor enter into chemical interactions with one another. The filling up of the lower branches of the split levels is accompanied by a gain in energy and these levels and the electrons which occupy them are called *bonding* levels and electrons. The filling up of the upper branches raises the energy of the system so these levels are said to be *antibonding*. The formation of stable molecules is possible if the total energy of the bonding electrons is greater than the total energy of the antibonding electrons.

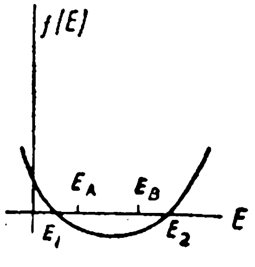


FIG. 44
Graphical solution of equation *3.30.*

Conditions favourable to the formation of strong bonds are

(1) the energies of the atomic levels, E_a and E_b, must not differ greatly,
(2) the wave functions ψ_a and ψ_b of the two atoms must overlap to the greatest possible extent,
(3) the functions ψ_a and ψ_b must have complete symmetry about the line of the bond.

Let us suppose that the energies E_a and E_b differ considerably. If $E_a \ll E_b$, then, as can be seen from [*35*], the two roots differ little from E_a and E_b respectively. Substituting $E = E_a$ in [*29*], we obtain $c_b = 0$, and for $E = E_b$ we find $c_a = 0$. This means that the molecular functions degenerate into the atomic functions ψ_a and ψ_b, and the atoms *a* and *b* do not form bonds with one another.

If the atomic functions ψ_a and ψ_b do not overlap, then, as can be seen from [*25*], the overlap integral $S = 0$, and also the resonance integral

$\beta = 0$. In this case we again obtain from [35] $E_1 = E_a$ and $E_2 = E_b$, which does not give a gain of energy when the atoms interact.

When the wave functions ψ_a and ψ_b overlap considerably, the integrals S and β may vanish if, in accordance with the conditions of symmetry for the atomic functions, each of them degenerates into two parts equal but with opposite signs. This takes place, for example, for $\psi_a = \psi_s$ and $\psi_b = \psi_{p_z}$, when the z-axis of atom b is perpendicular to the line of the bond (Fig. 45). In this case the overlap integral $\int \psi_a \psi_b \, d\tau$ breaks down into two integrals having the same magnitude, but opposite signs.

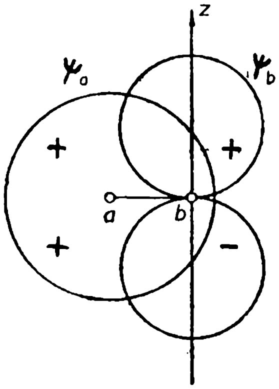


FIG. 45. Overlapping of ψ_s and ψ_{p_z} wave functions.

4. Forces of Repulsion

The Repulsive Potential according to the Classical Theory

If ions, atoms or molecules approach very closely, forces of repulsion begin to act between them. These forces are caused in the first place by the repulsion of the nuclei, in which considerable electric charges are concentrated. The magnitude of the forces of repulsion depends on the screening of the nuclei by the surrounding electrons and increases when the electron clouds of neighbouring atoms are superposed. Before the development of quantum mechanics, Bohr calculated the forces of repulsion from the static model of the atom. The arrangement of electrons in an 8-electron shell at the vertices of a cube led (Born-Landé) to forces of repulsion of a multipole character with a potential of the form b/r^n, where r is the distance between the nuclei of the atoms, and b and n are constants. This type of repulsive potential was widely used in calculating the energy of molecules and crystals, the values of the constants b and n being determined by experiment. The multipole model makes the forces of repulsion depend on the mutual orientation of the atoms, a result which is not confirmed by experiment and is not taken into account in the empirical power formula for the repulsive potential.

The Quantum Calculation of the Repulsive Potential

In the quantum mechanical calculation of the repulsive potential, the point distribution of the electrons is replaced by a distribution determined by the square of the modulus of the wave function, $(\psi)^2$, and leads to a fundamentally different formula for the repulsive potential.

Let us consider the problem of the repulsive potential in the simplest case, that of a proton penetrating a hydrogen atom. Let the nucleus of a hydrogen atom be situated at the point a (Fig. 46), and let the instantaneous position of the electron C be given by the vector r_a. We now bring a proton b close to the atom. The potential

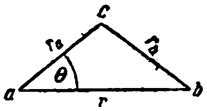


FIG. 46. Penetration of a proton into a hydrogen atom.

energy of the system consisting of the hydrogen atom and the proton will vary as

$$U = e^2\left(\frac{1}{r}-\frac{1}{r_b}\right) = e^2\left(\frac{1}{r}-\frac{1}{\sqrt{r^2+r_a^2-2rr_a\cos\theta}}\right). \qquad \ldots [3.36]$$

We find the mean value of the energy of interaction by averaging throughout the electron cloud

$$\bar{U} = (U\psi^2 dr). \qquad \ldots [3.37]$$

We take this energy as an energy of perturbation and use it to calculate the unperturbed atomic wave function. We introduce the variables

$$\xi = \frac{2r_a}{a_0}; \quad R = \frac{2r}{a_0} \qquad \ldots [3.38]$$

and expand U in powers of ξ/R:

$$\left.\begin{aligned} U &= -\frac{2e^2}{a_0}\left[\frac{\xi\cos\theta}{R^2}+\frac{\xi^2}{R^3}\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)+\ldots\right] \quad \text{for } \frac{\xi}{R}<1 \\ U &= -\frac{2e^2}{a_0}\left[\left(\frac{1}{\xi}-\frac{1}{R}\right)+\frac{R\cos\theta}{\xi^2}+\frac{R^2}{\xi^3}\left(\frac{3}{2}\cos^2\theta-\frac{1}{2}\right)+\ldots\right] \\ &\qquad \text{for } \frac{\xi}{R}>1. \end{aligned}\right\} \qquad \ldots [3.39]$$

For $\xi = R$ the two expressions become identical. The wave function for the ground state of hydrogen is

$$\psi(1s) = \left(\frac{1}{\pi a_0^3}\right)^{\frac{1}{2}} e^{-\frac{\xi}{2}}. \qquad \ldots [3.40]$$

When [37] is integrated, by using [39] and [40], all the terms containing spherical functions ($\cos\theta$, $3/2\cos^2\theta-1\ldots$) will vanish, apart from the first term in the second line of the expansion [39]. Hence

$$U_{\text{rep}} = -\frac{e^2}{a}\int\limits_{\xi=R}^{\infty}\left(\frac{1}{\xi}-\frac{1}{R}\right)e^{-\xi}\xi^2 d\xi = 2E_0\left(1+\frac{a_0}{r}\right)e^{-\frac{2r}{a_0}}, \qquad \ldots [3.41]$$

where E_0 is the energy of the ground state of hydrogen. The quantity U_{rep} is essentially positive, and increases rapidly as the distance diminishes. The rate at which U_{rep} increases is illustrated by the following figures:

r	$\frac{1}{2}a_0$	a_0	$2a_0$
U_{rep}	29·8	3·5	0·1 eV.

Formula [*41*], which was obtained by Unsold (1927), shows that the fundamental relationship between the repulsive potential and the distance is given by the exponential factor, which appears because of the exponential decay of the electron density in the atom. This relationship is characteristic of forces of repulsion as calculated according to quantum mechanics. The spherical symmetry of the repulsive potential in the case under consideration is due to the spherical symmetry of the function $\psi(s)$ and appears in all cases of filled electron subgroups.

Formula [*41*] can be generalised by considering the penetration of a proton into the K-shell of an atom with nuclear charge Z. For this case

$$U_{\text{rep}}(1s^2) = 4E_0 Z\left(1+\frac{a_0}{Zr}\right)e^{-\frac{2Z}{a_0}r}. \qquad \ldots [3.42]$$

When a proton penetrates into the L, M, N etc. shells of atoms, expressions of the form

$$U_{\text{rep}} = pE_0 Zf\left(\frac{r}{\rho}\right)e^{-\frac{r}{\rho}}, \qquad \ldots [3.43]$$

are obtained for the repulsive potential. They contain an exponential factor and, preceding it, a factor $f(r)$ of polynomial type, depending in a complicated manner on r. The fundamental relationship between U_{rep} and the distance is given by the exponential factor.

The Repulsive Potential for Atoms with Filled Shells

In this case also the repulsive potential contains an exponential factor which is dominant. For atoms of helium we have, according to Slater's calculations, at distances $r > 2a_0$

$$U_{\text{rep}}(\text{He}) = 481e^{-\frac{r}{0 \cdot 412}}. \qquad \ldots [3.44]$$

For atoms of Ne, Bleik and Mayer obtained

$$U_{\text{rep}}(\text{Ne}) = 1 \cdot 18 \times 10^4 e^{-\frac{r}{0 \cdot 395}}. \qquad \ldots [3.45]$$

For the repulsive potential of ions with filled shells Born and Mayer (1933) proposed the semi-empirical formula

$$U_{\text{rep}} = be^{-\frac{r}{\rho}}, \qquad \ldots [3.46]$$

which, like the power function br^{-n} introduced previously, contains two constants, b and ρ. The exponential form of the repulsive energy which emerges from the quantum calculation, is in better agreement with the experimental data. Thus, the empirical value of n determined for alkali-halogen salts varies within broad limits from $n = 5$ for lithium salts to $n = 11$ for caesium salts, while the empirical values of ρ obtained for 17 alkali halides give a mean deviation of 6% from ρ_{mean}, $= 0 \cdot 345$.

The rate at which the repulsive energy increases as the atoms approach one another and the electron clouds overlap enables us to understand the impenetrability and hardness which are the fundamental properties of the solid state. The steeply rising branch of the curve of U_{rep} as $r \to 0$ enables us to introduce an important characteristic of atoms—the atomic radii. The equilibrium distance between the particles depends also on the type of attractive forces, and so there are differences between atomic, ionic and molecular radii (*vide* Table 37, pp. 192-3).

The Total Potential Energy of Interaction of Two Particles

The energy of interaction may be expressed approximately as the sum of two terms,

$$U = U_{att} + U_{rep}, \qquad \ldots [3.47]$$

corresponding to the potentials of the forces of attraction and repulsion. Fig. 47*a* shows curves of these potentials, and Fig. 47*b* the curve of their sum. The position of

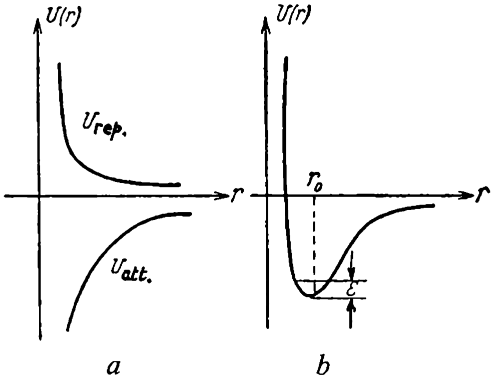


FIG. 47. (*a*) Attractive and repulsive potentials, (*b*) total potential energy of interaction of two particles.

the minimum on this curve gives r_0, the equilibrium distance between the particles, and the depth U_0 of the minimum is the binding energy of the particles. When the zero point energy $\varepsilon = \frac{1}{2}h\nu$ is added, the binding energy is reduced and becomes $U_0 - \varepsilon$. For particles with different masses m_1 and m_2 (isotopes) the frequencies of oscillation, ν_1 and ν_2, are different [*1.130*]. A lighter isotope corresponds to a higher zero point energy.

5. Molecular Forces

Dispersion Forces

The existence of the liquid and solid phases in inert elements with low boiling and melting points shows that weak forces of attraction operate even between atoms with filled shells. Analogous forces operate between molecules with saturated bonds. In both cases these forces are called van der Waals forces. They are the reason for the first correction needed in Clapeyron's equation $pv = RT$ and its replacement by van der Waals' equation $(p + a/v^2)(v - b) = RT$. The second correction in this equation takes account of the forces of repulsion.

Atoms with filled valence shells have a spherical distribution of electric charge. They have thus no constant electric moments and they do not excite induced moments in each other. The van der Waals forces are due to the presence of significant *instantaneous* dipoles in the atoms. Thus in the hydrogen atom the mean electric moment is zero, while the instantaneous moment is a significant quantity $p = ea_0 = 4{\cdot}8 \times 0{\cdot}53 \times 10^{-18}$ c.g.s.u. $= 2{\cdot}5$ debye (1 debye $= 10^{-18}$ c.g.s. units of electric moment). When the atoms approach to within a distance such that there is as yet no overlapping of their wave functions, the instantaneous atomic dipoles interact. The energy of this interaction may be reduced if the electric moments become arranged in a regular manner. This is brought about by harmonisation (synchronisation) of the movements of electrons in neighbouring atoms, and gives rise to forces of attraction. The quantum mechanical calculation of these forces was given by London (1930).

As the simplest model, we shall consider two identical electrical oscillators in which the positive charges are fixed, while the negative charges $-e$ oscillate along the x-axis about the equilibrium positions. The dipole moment of the oscillator is

$$\mu = ex = \alpha E, \qquad \ldots [3.48]$$

where x is the displacement of the negative charge, α is the polarisability of the dipole and E is the intensity of the field. For a quasi-elastic oscillator

$$eE = Kx, \qquad \ldots [3.49]$$

where K is the elastic modulus. From [*48*] and [*49*] we find

$$K = \frac{eE}{x} = \frac{e^2}{\alpha}. \qquad \ldots [3.50]$$

The frequency of vibration of the oscillators is

$$\nu = \frac{1}{2\pi}\sqrt{\frac{e^2}{\alpha m}}. \qquad \ldots [3.51]$$

In the absence of interaction between the oscillators their total energies are

$$\mathscr{E}_1 = \frac{p_1^2}{2m}+\frac{e^2x_1}{2\alpha} \quad \text{and} \quad \mathscr{E}_2 = \frac{p_1^2}{2m}+\frac{e^2x_2}{2\alpha}. \qquad \ldots [3.52]$$

Here x_1 and x_2 are the displacements, and p_1 and p_2 the momenta, of the oscillators.

We now place the oscillators on the x-axis at a distance R from one another and calculate the energy of electrostatic interaction between them:

$$\frac{e^2}{R}+\frac{e^2}{R+x_1+x_2}-\frac{e^2}{R+x_1}-\frac{e^2}{R+x_2} = -\frac{2e^2x_1x_2}{R^3}+\ldots \qquad \ldots [3.53]$$

In the L.H.S. of [*53*] the first two terms express the energies of repulsion of the positive and negative charges, and the two other terms the energies of attraction of opposite charges. If we suppose that R is large by comparison with the displacements x_1 and x_2, expanding the energy E in a series and taking only the first term of the expansion, we obtain the expression on the R.H.S. The total energy of the system, allowing for the interaction, will be

$$\mathscr{E} = \frac{1}{2m}(p_1^2+p_2^2)+\frac{e^2}{2\alpha}(x_1^2+x_2^2)-\frac{2e^2x_1x_2}{R^3}. \qquad \ldots [3.54]$$

The interaction of the oscillators removes the degeneracy and causes their frequencies to split up. To calculate this splitting, we introduce new (normal) coordinates

$$x_s = \frac{1}{\sqrt{2}}(x_1+x_2); \quad x_a = \frac{1}{\sqrt{2}}(x_1-x_2). \qquad \ldots [3.55]$$

In these coordinates we obtain for the momenta ($p = m\dot{x}$) the expressions

$$p_s = \frac{1}{\sqrt{2}}(p_1+p_2); \quad p_a = \frac{1}{\sqrt{2}}(p_1-p_2), \qquad \ldots [3.56]$$

and for the energy of the system

$$\mathscr{E} = \frac{1}{2m}(p_s^2+p_a^2)+\frac{e^2}{2\alpha}(x_s^2+x_a^2)+\frac{e^2}{R^3}(x_s^2-x_a^2). \qquad \ldots [3.57]$$

This latter expression may be transformed into

$$\mathscr{E} = \left[\frac{1}{2m}p_s^2+\left(\frac{e^2}{2\alpha}+\frac{e^2}{R^3}\right)x_s^2\right]+\left[\frac{1}{2m}p_a^2+\left(\frac{e^2}{2\alpha}-\frac{e^2}{R^3}\right)x_a^2\right], \qquad \ldots [3.58]$$

corresponding to the energy of two non-interacting oscillators, with eigenfrequencies equal to

$$\nu_s = \frac{1}{2\pi}\sqrt{\frac{e^2}{m\alpha}\left(1+\frac{2\alpha}{R^3}\right)}; \quad \nu_a = \frac{1}{2\pi}\sqrt{\frac{e^2}{m\alpha}\left(1-\frac{2\alpha}{R^3}\right)}. \qquad \ldots [3.59]$$

The total zero point energy of these oscillators is

$$\varepsilon = \tfrac{1}{2}h\nu_s+\tfrac{1}{2}h\nu_a = \tfrac{1}{2}h\nu\left(\sqrt{1+\frac{2\alpha}{R^3}}+\sqrt{1-\frac{2\alpha}{R^3}}\right). \qquad \ldots [3.60]$$

Expanding this in a series, we obtain

$$\varepsilon = h\nu\left(1-\frac{\alpha^2}{2R^6}+\ldots\right). \qquad \ldots [3.61]$$

Formula [*61*] shows that the total zero point energy $h\nu$ of the oscillators is diminished as a result of the interaction in inverse proportion to the sixth power of the distance between the oscillators. This corresponds to the appearance of a force of attraction between them the magnitude of which depends on the polarisability of the atoms (oscillators), which in turn determines their optical properties (dispersion of light). For this reason the forces mentioned are known as *dispersion forces.*

For the simplest atoms the potential of the dispersion forces can be calculated by taking into account higher terms in the multipole interaction. For hydrogen atoms an approximate formula (ignoring the zero point energy) can be obtained:

$$U_{\text{disp.}} = -6\frac{e^2}{a_0}\left(\frac{1}{r^6}+\frac{22{\cdot}5}{r^8}+\frac{236}{r^{10}}\right). \qquad \ldots [3.62]$$

At the distance at which the van der Waals forces operate, the second term in [*62*] is about 10% as large as the first, while the third term is insignificantly small. For atoms of helium

$$U_{\text{disp.}} = -1{\cdot}62\frac{e^2}{a_0}\left(\frac{1}{r^6}+\frac{7{\cdot}9}{r^8}+\frac{30}{r^{10}}\right). \qquad \ldots [3.63]$$

When the repulsive potential is taken into account, the curve of the resulting interaction has a slight minimum at $r = 5{\cdot}5a_0$ with a total depth of $0{\cdot}75 \times 10^{-3}$ eV, which explains the low binding energy of liquid helium. At normal pressures helium remains fluid right down to 0° K, since in consequence of the feebleness of the forces of attraction crystalline helium 'melts' on account of its own zero point energy. Crystalline helium is formed only under pressures above 25 atm.

Dispersion forces are universal forces, operating not only between atoms with filled shells, but also between any atoms, ions and molecules. In the presence of a strong interaction (the chemical bond) the dispersion interaction is only a minor correction. In other cases the dispersion interaction forms a significant part of the general intermolecular interaction, while in cases analogous to the interaction of atoms of the inert elements, the dispersion forces are the only attractive forces between the particles.

Interaction of Molecules with Constant and Induced Dipoles

If molecules possess constant dipole moments, as do for example, molecules of water, or if when their polarisability is high they acquire induced dipoles, then an additional dipole interaction appears. The energy of this interaction is electrostatic and, like the energy of the dispersion interaction, diminishes in inverse proportion to the sixth power of the distance.

If the molecules possess constant dipole moment μ, then under the influence of electric forces they tend to orient themselves in relation to one another in such a way as to decrease their energy of interaction (Fig. 48*a*). This orientation is disturbed by random thermal agitation (Fig. 48*b*). The electrostatic interaction energy depends to a high degree on the temperature. At sufficiently high temperatures, when the interaction energy of two dipoles $U_{\text{dip}} \ll kT$, the orientational interaction energy was shown by Keesom to be

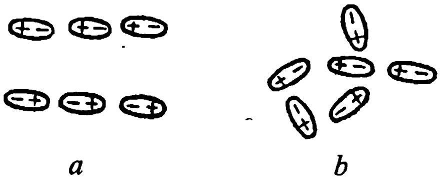


FIG. 48. Mutual orientation of dipoles: (*a*) $T = 0$, (*b*) $T > 0$.

$$U_{\text{or}} = -\frac{2}{3}\frac{p^4}{r^6}\cdot\frac{1}{kT}. \qquad \ldots [3.64]$$

At low temperatures when $U_{\text{dip}} \gg kT$, if the dipoles are perfectly orientated, the dipole interaction energy is

$$U_{\text{or}} = -\frac{2p^2}{r^3} \quad (T \to 0).$$

In molecules which have appreciable polarisability α, induced dipole moments $p = \alpha E$ appear, where E is the electric field intensity. The interaction energy of the induced dipoles does not depend on the temperature. As Debye showed,

$$U_{\text{ind}} = -\frac{2\alpha p^2}{r^6}. \qquad \ldots [3.65]$$

In the general case the interaction energy of two molecules can be expressed as a sum of different parts corresponding to the orientation, induction and dispersion interactions:

$$U_{\text{mol}} = -\frac{1}{r^6}\left(\frac{2}{3}\frac{p_1^2 p_2^2}{kT}+\alpha_2 p_1^2+\alpha_1 p_2^2+\frac{3h}{2}\alpha_1\alpha_2\frac{\nu_1\nu_2}{\nu_1+\nu_2}\right). \qquad \ldots [3.66]$$

In different molecules the part played by each of these terms is different (Table 16).

Table 16

Types of Intermolecular Interaction (%)

Molecule	U_{or} (orientation)	U_{ind} (induction)	U_{disp} (dispersion)
CO	0	0	100
H_2O	77	4	19
CH_3OH (methyl alcohol)	64	14	22

The Hydrogen Bond

Bonds between a hydrogen atom and the more electronegative atoms F, O, and N have a partly ionic character. Thus the molecules HF, H_2O and NH_3 have constant dipole moments; between them and others similar, considerable dipole interaction is possible leading to the formation of associations, and at appropriate temperatures even crystals. Fig. 49 shows in schematic form the formation of intermolecular hydrogen bonds between molecules HF, H_2O and NH_3. The hydrogen atoms in these molecules form sufficiently strong covalent-ionic bonds with intramolecular distances E–H of 0·9, 1·07 and 1·09 Å respectively. The dipole moment of the O–H bond is $1{\cdot}51\times10^{-18}$ c.g.s. units and of the NH bond $1{\cdot}31\times10^{-18}$ c.g.s. units. Interaction of the dipoles causes neighbouring molecules to orient themselves in such a way that the most electropositive parts of the molecules, that is the hydrogen atoms, approach the electronegative atoms F, O and N. The shortest distances between neighbouring molecules, which are shown in Fig. 49 by dotted lines representing intermolecular hydrogen bonds, depend on the mutual orientation of the molecules. The means of the distances H . . . F, H . . . O and H . . . N are equal to 1·6, 1·8 and 1·9 Å, that is, significantly more than the intramolecular distances. The normal

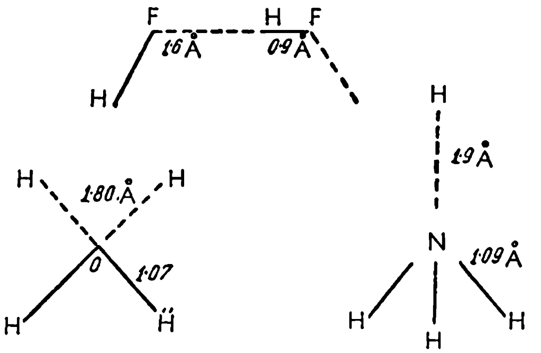


Fig. 49. Formation of hydrogen bonds.

number of hydrogen bonds formed by the molecules HF and NH_3 is 1, but for a water molecule it is 2. Water and ammonia molecules have a characteristic tetrahedral distribution of intra- and intermolecular bonds. The energy of the intermolecular interaction known as the *hydrogen bond* is small ($\sim$5 kcal./mole). Nevertheless the hydrogen bond is very important. The reasons are, firstly that it is so common and secondly, that because the energy is not too large the hydrogen bond is fundamental in processes which take place at normal temperatures, in particular in biological processes. The hydrogen bond is the cause of the unusual electrical properties of the ferroelectric salt $NaKC_4H_4O_6$, $4H_2O$, KH_2PO_4 and other analogous ferroelectrics. Some surprising properties of water are due to the hydrogen bond. The melting and boiling points of molecular substances in the series H_2Te, H_2Se, H_2S, which are determined by the van der Waals forces, drop as the molecular weight decreases (Fig. 50). If these curves are extrapolated, then for water we obtain $t_{\text{boiling}} = -80°$ C. and $t_{\text{melting}} = -100°$ C., whereas the experimental values lie considerably higher. The rise in the boiling and melting points of HF, H_2O and NH_3 is due to the hydrogen bonds, which almost double the boiling point of water on the absolute scale. The effect of the hydrogen bonds in the case of water is significantly greater than for HF and NH_3. The explanation for this is that each water molecule forms twice as many hydrogen bonds as the molecules of HF and NH_3. The methane molecule CH_4 does not form hydrogen bonds, and so the melting and boiling points in the series SnH_4, GeH_4, SiH_4 and CH_4 vary monotonically.

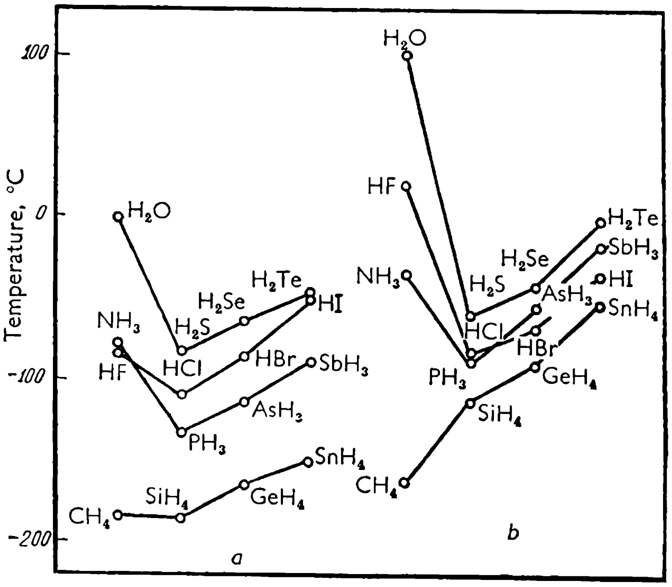


FIG. 50. Anomalies in melting (*a*) and boiling (*b*) points of hydrogen bonded molecular hydrides.

Problems

1. Find the first excitation potential $\mathscr{E}_1$ of the hydrogen atom as a function of the ground state energy I.

2. Given that the first excitation potential of the oxygen molecule is $\mathscr{E}_1 = 2{\cdot}6$ eV, find: (*a*) at what velocity of collision its impact with a wall will be inelastic; (*b*) at what temperature the mean thermal energy of the molecule is sufficient for thermal excitation.
3. What is the explanation for the decrease in valency in the series of Sn and Pb compounds to 2, and for Tl to 1, compared with normal valencies of 4 and 3 for elements of the IVB- and IIIB-subgroups?

Chapter 4

THE BASIC TYPES OF CHEMICAL BOND

1. The Covalent Bond

The Molecular Ion of Hydrogen H_2^+

The ionised molecule of hydrogen H_2^+ contains one electron. This molecular ion occupies in the quantum theory of molecules a position analogous to that occupied by the hydrogen atom in the theory of the atom. Let us consider the solution of the problem for this ion. The Hamiltonian of the H_2^+ ion, excluding the term corresponding to the nuclear repulsion energy, takes the form

$$H = -\frac{h^2}{8\pi^2 m}\nabla^2_{xyz} - \frac{e^2}{r_A} - \frac{e^2}{r_B}, \qquad \ldots [4.1]$$

where r_A and r_B are the distances of the electron from the two nuclei. The atomic functions $\psi_A = \psi_B$ satisfy Schrödinger's equation for the hydrogen atom, $H\psi_A = E_0\psi_A$, or in expanded form

$$\left(-\frac{h^2}{8\pi^2 m}\nabla^2 - \frac{e^2}{r_A}\right)\psi_A = E_0\psi_A, \qquad \ldots [4.2]$$

where $E_0 = -e^2/2a_0$ is the energy of the ground state of the atom. From [1] and [2] $H\psi_A = E_0\psi_A - (e^2/r_B)\,\psi_A$, consequently

$$E_A = \int \psi_A H \psi_A d\tau = E_0 - \int \frac{e^2}{r_B}\psi_A^2 d\tau, \qquad \ldots [4.3]$$

$$\beta = \int \psi_B H \psi_A d\tau = E_0 S - \int \frac{e^2}{r_B}\psi_A\psi_B d\tau. \qquad \ldots [4.4]$$

Substituting [3] and [4] in [3.34], we obtain the electronic part of the energy of the ion

$$E_{1,2} = E_0 - \frac{\int \left(\frac{e^2}{r_B}\right)\psi_A^2 d\tau \pm \int \left(\frac{e^2}{r_B}\right)\psi_A\psi_B d\tau}{1 \pm \int \psi_A\psi_B d\tau}. \qquad \ldots [4.5]$$

Using the wave functions of the hydrogen atom one can find the values of the integrals in [5] for different values of the internuclear distance. The results of calculating the electron energy (curve *a*), the nuclear repulsion energy (curve *b*) and the total energy (curve *c*) for the states $\psi_A + \psi_B$ and $\psi_A - \psi_B$ are given in Fig. 51. The curve of the

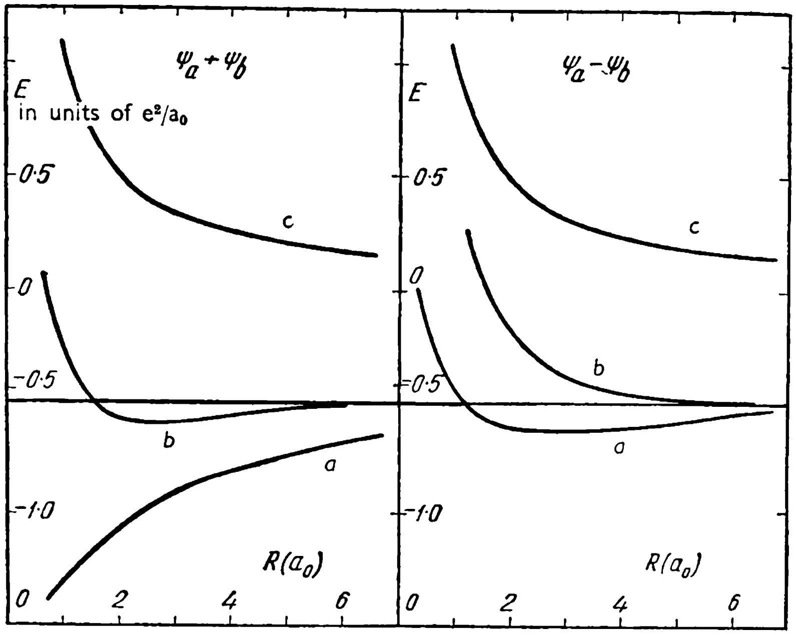
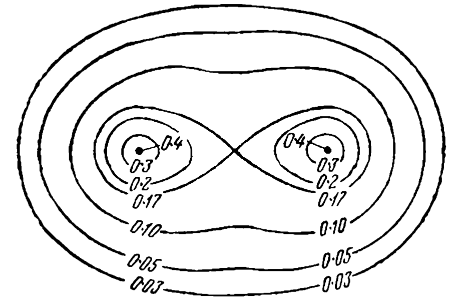
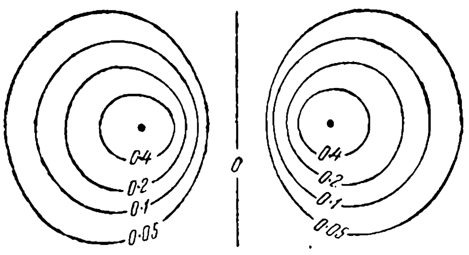


FIG. 51. The electron repulsion energy (*a*), the nuclear repulsion energy (*b*), and the total energy of the ion H_2^+ (*c*) as functions of the internuclear distance.

total energy for the first state has a minimum at the distance $r = 2a_0 = 1{\cdot}06$ Å, which corresponds to the equilibrium distance of the ion H_2^+ and agrees well with experiment. The curve of the total energy of the second state has no minimum. In Fig. 52 the electron density distribution is shown in terms of contours of equal ψ^2 for both states ($\sigma 1s$ and $\sigma^* 1s$—stable and excited). Fig. 53 shows the charge distribution along the axis of the molecule and for comparison it also gives the curves (dotted lines) of electron density for the free atoms, ψ_A^2 and ψ_B^2. The additive curve for H_2^+ is their average $\frac{1}{2}(\psi_A^2+\psi_B^2)$. Comparison of the curves shows that in the ground state of the molecule there is a considerable increase in the electron density between the nuclei. Because of this, strong forces of attraction arise and a stable molecule is formed. In the excited state the electron charge density between the nuclei falls to zero. The nuclei are then less effectively screened, and the forces of repulsion between them exceed the forces of attraction. As a result the atoms are repelled from one another and no molecule is formed.



σ-1s (attractive)



σ-1s (repulsive)

FIG. 52. Electron density distribution in the H^+ ion.

[Figs 51, 52 and 53 are reproduced from *Valence* by C. A. Coulson by kind permission of Professor Coulson and the Clarendon Press, Oxford.]

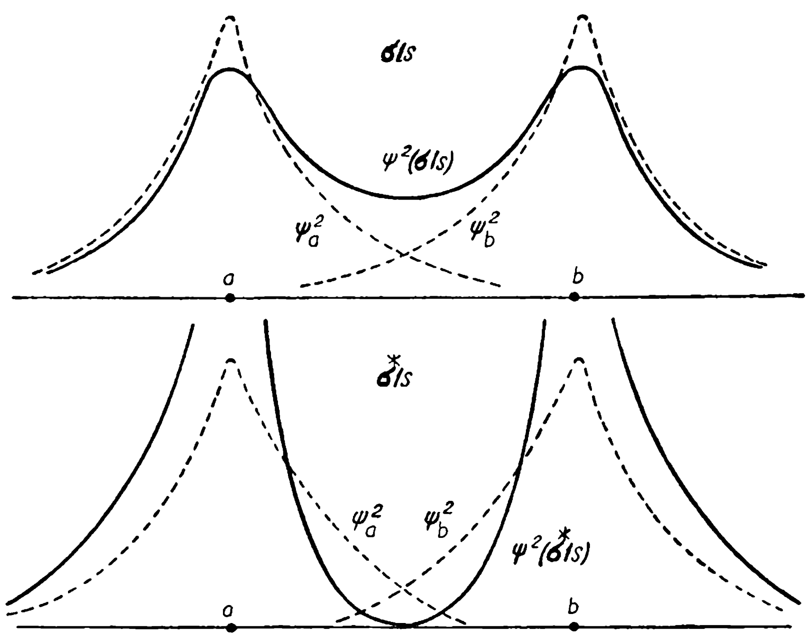


FIG. 53. Electron density distribution along the axis of the ion.

As Fig. 52 shows, formation of the H_2^+ ion involves considerable redistribution of the electron density as compared with the isolated atoms. The use of the unperturbed atomic functions ψ_a and ψ_b to find the molecular functions introduces a certain error, which may be removed by a suitable choice of perturbed functions. For example, we can take the atomic functions of hydrogen with altered effective nuclear charge Z, i.e.

$$\psi_a = \left(\frac{Z^3}{\pi a_0^3}\right)^{\frac{1}{2}} e^{-\frac{Zr_a}{a_0}}. \qquad \ldots [4.6]$$

The value of Z is found by a variational method, which involves determining the minimum energy E for a given R. Using the observed equilibrium distance R = 1·06 Å of the H_2^+ ion, we obtain an effective charge value Z = 1·24, which considerably improves the theoretical value for the energy of the ion. The next approximation might be to take account of the polarisation of the electron clouds of the atoms in the molecule, that is to replace the spherically symmetric by axially symmetric functions, for example of the type

$$\psi_a = e^{-\frac{Zr_a}{a_0}} + cxe^{-\frac{Zr_a}{a_0}} \qquad \ldots [4.7]$$

(the x-axis coincides with the axis of the molecule). The parameters c and Z are determined from the energy minimum. This gives a more accurate value for the energy of the molecule, making the theoretical value coincide very closely with experiment.

The ellipsoidal form of the electron cloud in the dumb-bell molecule suggests the possibility of an empirical choice of wave functions. James showed that a very good electron density distribution and a more accurate energy value for H_2^+ are obtained if the molecular wave function is taken in the form

$$\psi_1 = e^{-c_1(r_a+r_b)}\{1+c_2(r_a-r_b)^2\}, \qquad \ldots [4.8]$$

where the parameters c_1 and c_2 are found from the minimum energy.

The Hydrogen Molecule H_2

This, the simplest molecule, differs from the H_2^+ ion in that it contains two electrons. The solution of the problem of the hydrogen molecule by the method of molecular orbitals is analogous to the solution of the problem of the H_2^+ ion. The ground state of the molecule corresponds to an arrangement of the two electrons on the lowest energy level with opposite spin orientations. The electron charge density in between the nuclei in the hydrogen molecule is higher than the charge density in the molecular ion. This necessarily leads to an increase in the energy of the molecule and a decrease in the internuclear distance, which is indeed confirmed by experiment. The equilibrium distance between the nuclei in the H_2 molecule ($R = 1{\cdot}38\ a_0 = 0{\cdot}74$ Å) is considerably less than in the ion H_2^+ ($R = 2{\cdot}00\ a_0 = 1{\cdot}06$ Å). The dissociation energy of the H_2 molecule (4·72 eV) is almost twice the dissociation energy of the H_2^+ ion (2·64 eV). In the energy balance of the molecule the Coulomb nuclear repulsion term is much greater due to the nuclei being nearer (19·3 eV in H_2 and 13·5 eV in H_2^+), and a new term appears, corresponding to the repulsive energy of the electrons (17·8 eV). The dissociation energy of the molecule is equal to the difference between the energy of the molecule and the total energy of two hydrogen atoms, the latter being twice the potential energy of a hydrogen atom ($2\times 13{\cdot}6 = 27{\cdot}2$ eV)

$$D_{H_2} = -(2\times 34{\cdot}0)+17{\cdot}8+19{\cdot}3-(-27{\cdot}2) = -3{\cdot}7\ \text{eV}. \qquad \ldots [4.9]$$

The first term in [*9*] is equal to twice the energy of each electron in the molecular orbital (−34·0 eV), and the second and third terms are the repulsive energies of the electrons and nuclei. The dissociation energy is thus the difference of two large quantities; in this approximation the value obtained is 1·0 eV less than the experimental value, which corresponds to the mean error in the self-consistent field method as applied to molecules. On taking into account the various corrections, we can obtain a theoretical value for the dissociation energy of the hydrogen molecule which agrees with the experimental value.

The Lithium Molecule Li_2

Spectroscopic data enable us to establish the existence of diatomic Li_2 molecules and to determine their dissociation energy (1·14 eV) and internuclear distance (2·67 Å). The calculation for the Li_2 molecule is carried out in a similar way to that for the hydrogen molecule, but it has to be remembered that the valence electrons are 2*s*, while the 1*s* atomic levels are filled. The dissociation energy obtained in this way (1·09 eV) agrees well with experiment. The increased internuclear distance in the Li_2 molecule, is a result of the increased repulsive forces and makes its dissociation energy significantly less than that of the hydrogen molecule.

Single and Multiple Bonds, σ and π-bonds

In finding molecular orbitals by the method of linear combinations $\psi_A \pm \psi_B$, one can choose as initial functions ψ_A and ψ_B not only the 1*s*-functions of the atoms *A* and B, but also other atomic $2s$-, $2p_x$-functions etc. The valence electron density distribution in the molecule depends on the type of functions chosen. Fig. 54 shows the contours and signs of the fields of the molecular wave functions formed by superimposing the s- or p-functions. The presence or absence of a centre of symmetry in the wave functions of the molecules is indicated by the signs g or u, which show whether the functions are even or odd. The molecular wave functions formed by

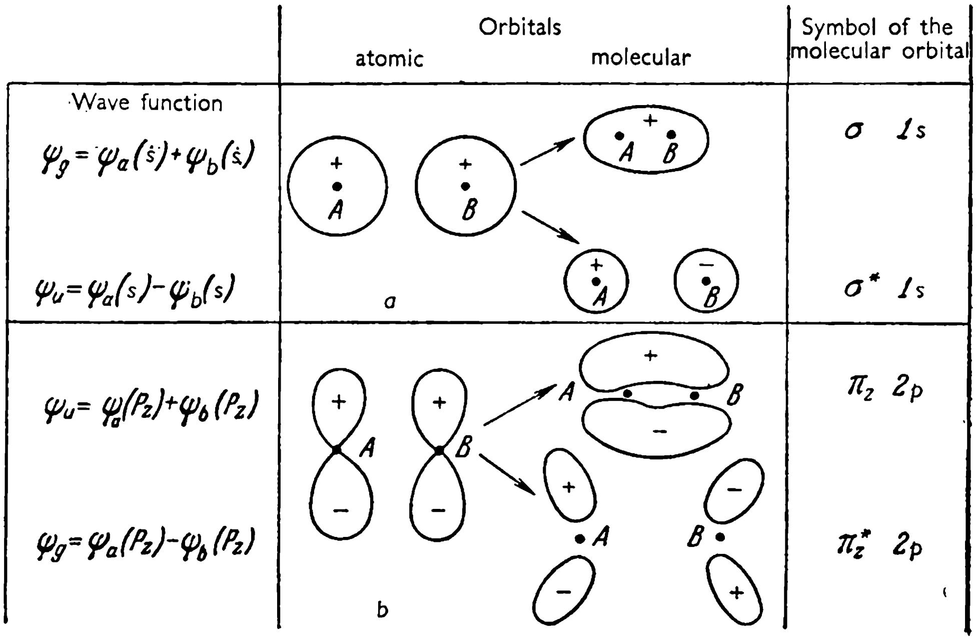

Orbitals
atomic
molecular
Symbol of the molecular orbital
Wave function
$\psi_g = \psi_a(s) + \psi_b(s)$
$\psi_u = \psi_a(s) - \psi_b(s)$
$\psi_u = \psi_a(p_z) + \psi_b(p_z)$
$\psi_g = \psi_a(p_z) - \psi_b(p_z)$
a
b
A
B
+
−
$\sigma\ 1s$
$\sigma^*\ 1s$
$\pi_z\ 2p$
$\pi_z^*\ 2p$

FIG. 54. Molecular wave functions formed from s- and p-functions.

two s- or two p_x-functions possess axial symmetry (Fig. 54). The valence electron density in this case is a maximum on the axis of the molecule.

An essentially different electron density distribution is obtained when two p_y- or two p_z-electrons interact. The clouds of these electrons have the form of dumb-bells whose axes are directed perpendicularly to the axis of the molecule. The linear combination $\psi_A(p_z)+\psi_B(p_z)$ corresponds to condensation of the electron density and describes the stable state of the molecule. The molecular wave function does not have a centre of symmetry. The electron density distribution is characterised by two regions extending in a direction parallel to the axis of the molecule (Fig. 54). Between these regions passes the nodal plane of the wave function, which contains the axis of the molecule and the z-axis, and for which $\psi = 0$. The electron density distribution does not have axial symmetry. This type of bond is formed by two p_y electrons. The electron density maxima in this case lie in a plane turned through 90° and containing the y-axis. An axially symmetric electron density distribution arises if the two bonds, formed by p_y- and p_z-electrons respectively, are superimposed.

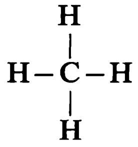
Bonds for which the electron density is concentrated along or perpendicular to the axis of the molecule are called σ- and π-bonds respectively. The difference in the electron density distribution results in a difference in the bond energy. The overlapping of the wave functions for the σ-bonds (other conditions being the same) is greater than for the π-bonds, so the energy of the σ-bonds exceeds the energy of the π-bonds. For carbon atoms at a distance of 1·54 Å the ratio of the overlap integrals for the $p\sigma$- and $p\pi$-bonds is 2:1. When atoms interact, σ-bonds are formed first, and then one or two π-bonds in addition, depending on the number of valence electrons. The corresponding bonds are known as double or triple bonds; simple σ-bonds are called single bonds.

Each energy level in a molecule, just as in an atom, can by Pauli's principle accommodate two electrons with opposite spins. Two electrons take part in a single bond, four in a double bond, and six in a triple bond. The fact that the energy of the π-bonds is less than that of the σ-bonds explains the increased reactivity of compounds with multiple bonds compared to compounds with single bonds. In the chemistry of carbon, multiple bonds are very common and are the important ones in aromatic and in general in the cyclic compounds of carbon, where the π-electron interaction is the reason for the plane form of the molecule (benzene, naphthalene, plane atomic lattices in graphite). Aliphatic compounds of carbon are characterised by single bonds.†

Multiplicity of Bonds and Interatomic Distances

An atom with valency v can form a different number k of bonds. If the atom forms only single bonds, its valency is the same as the number of possible bonds: $v = k$. Single bonds are formed in compounds of a multivalent atom with univalent atoms and in these compounds, for example in molecules of the hydrides HF, H_2O, NH_3, CH_4, it is easy to establish the valency of the second atom. The case $k<v$ is the case of multiple bonds. In covalent compounds the multiplicity $\kappa = v/k$ of the bonds varies within the limits $1 \leqq \kappa \leqq 3$. The case $k>v$, for which $\kappa<1$, relates to the metallic and ionic bond. The higher the multiplicity of the bond, represented in chemistry by the number of valence lines, the greater the interaction energy of the atoms and the shorter the interatomic distances. This may be illustrated by the example of the bonds formed by carbon.

In molecules of methane, carbon dioxide and prussic acid, which have the structures

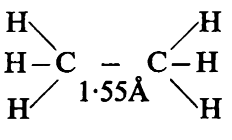


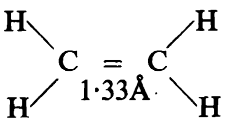
$$\begin{array}{c} H \\ | \\ H-C-H \\ | \\ H \end{array}$$

tetrahedron

$O=C=O$ $\qquad$ $H-C\equiv N$,

the tetravalent carbon atom is capable of forming single, double and triple bonds. In molecules of ethane, ethylene and acetylene, which have the structures





$H-C\equiv C-H$,
1·20Å

the carbon atoms form multiple bonds with one another. The interatomic distances C–C in these molecules diminish regularly with increasing bond multiplicity. In diamond and graphite crystals the carbon atoms form 4 and 3 bonds respectively with multiplicities $\kappa = 1$ and $\kappa = 4/3 = 1{\cdot}33$. The difference in the multiplicities corresponds to the difference between the interatomic distances C–C in diamond (1·54 Å) and in graphite (1·41 Å). In the benzene molecule C_6H_6, which has the form of a regular hexagon, each carbon atom forms two bonds with carbon atoms and one bond with a hydrogen atom. The multiplicity of the C-C bonds in benzene; $\kappa = (4-1)/2 = 3/2 = 1{\cdot}50$, is higher than in graphite, but the interatomic distance 1·39 Å is slightly less. In Fig. 55 the length of the C–C bond is shown as a function of the multiplicity. An increase of the multiplicity from $\kappa = 1$ to $\kappa = 3$ corresponds to a shortening of the bond by 0·34 Å, that is, by 22% compared with a single bond. With transition metals of the IVA- and VA subgroups carbon forms carbides with a cubic structure of the NaCl type and coordination number $k = 6$. If we assume that in these carbides, carbon keeps its valency $v = 4$, we find that the multiplicity of its bonds is $\kappa = 4/6 = 2/3$. From the length of the M–C bond, using a value for the atomic radii of the metals corrected to coordination number $k = 6$, we obtain for the radius of the carbon atom in these carbides the value 0·80 Å, which exceeds the covalent radius of carbon for the single bond (0·77 A). The value calculated from this for the length of the C–C bond with $\kappa = 2/3$ is marked in Fig. 55 by a broken line. The fact that the valency of an atom is conserved in different compounds independently of the coordination number shows that the multiplicity of the bond may be either integral or fractional.

---

† With increasing interatomic distance the strength of the π-bonds diminishes more rapidly than that of the σ-bonds, so multiple bonds are more readily formed by atoms of small radius. Apart from carbon, multiple bonds are formed only by nitrogen, oxygen and, in special cases, sulphur.

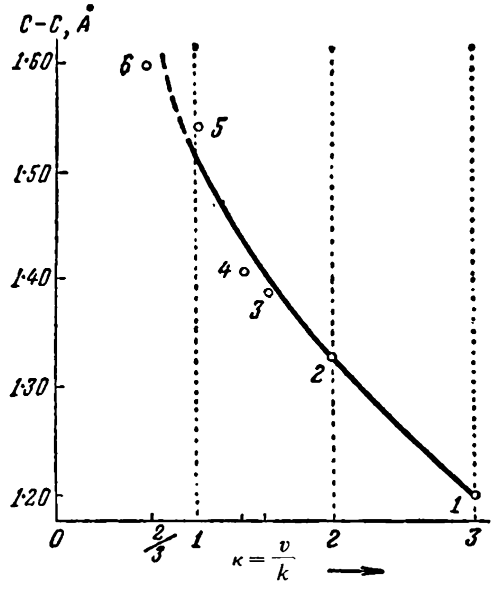


FIG. 55. Dependence of the inter-atomic distance C–C on the multiplicity of the bond: 1, acetylene; 2, ethylene; 3, benzine; 4, graphite; 5, diamond, ethane; 6, metallic carbides.

For metallic and ionic bonds $\kappa < 1$. In these cases too, the interatomic distances depend on the multiplicities of the bonds. This relationship is usually expressed in the form of a correction to the atomic and ionic radii for different coordination numbers (*vide* Table 37, *a* and *b*, pp. 192-3). Variations in the coordination number are not always caused by variations in the multiplicities of the bonds. In some cases they are due to variations in the valency of the atom. Thus in the carbon monoxide molecule CO and in the carbon dioxide molecule CO_2 (which has a linear symmetric structure O=C=O) the interatomic oxygen-carbon distances are similar (1·13 and 1·15 Å) and correspond to the length of the double bond C=O. The lowering of the coordination number of the carbon atom in the CO molecule is due to a lowering of its valency. The incomplete use of the valency of the carbon atom is shown by the increased reactivity of carbon monoxide compared to the chemically inert carbon dioxide.

Level Schemes for Diatomic Molecules

The total scheme of levels for a molecule consists of the set of all the split atomic levels. The amount of the splitting depends on the degree of overlapping of the wave functions, which, other things being equal, increases with diminution of the internuclear distance. It is essential that the scheme of molecular levels should transform into the scheme of atomic levels not only as $R \to \infty$ (initial atoms A, B), but also when $R \to 0$. This corresponds to merging of the two nuclei and the formation of a 'united atom' A+B.

Fig. 56 shows the scheme for molecules consisting of identical atoms. The lowest level $\sigma 1s$ of the molecule transforms into the lowest level $1s$ of the united atom. The next level, $\sigma^* 1s$ transforms into the $2p$ level, because of resemblance between the symmetries of the corresponding wave functions (Fig. 54). Fig. 57 represents the sequence of lowest energy levels of diatomic molecules and, by way of an example, the electron structure of the O_2 molecule. Knowing the sequence of levels, one can find the electron structure of the molecule and determine a number of its properties (bond strength, magnetic properties etc.). The position of the energy levels, as may be deduced from Fig. 56, varies according to the internuclear distance. So neighbouring levels, for example $\sigma 2p$ and $\pi 2p$, and especially levels with higher principal quantum numbers, may overlap. Table 17 shows the electron structure of some of the simplest molecules and their ions, and gives the multiplicity of the bond (the single bond being taken as unity), calculated as the difference between the numbers of *bonding and anti-bonding* electrons (that is electrons on molecular levels lying respectively lower and

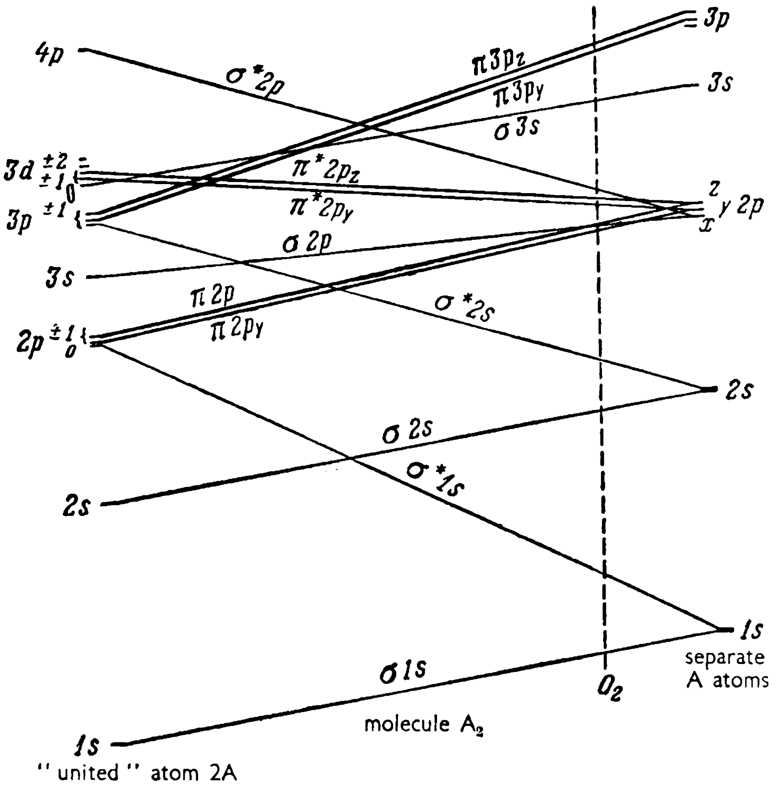


Fig. 56. Scheme of levels of diatomic molecules of type A_2.

higher than the unsplit atomic levels); and it also gives the dissociation energies. The last column shows the spectroscopic symbols for the states of the molecule, in which the upper index stands for the multiplicity of the spectrum (1—singlets, 2—doublets, 3—triplets), the Greek letter denotes the total moment of the electrons about the axis, and the subscript gives the symmetry of the molecule.

The data of Table 17 show that, in agreement with experiment, the inert gases, and also beryllium, do not form diatomic molecules. Atoms of the other elements form molecules whose dissociation energies correspond to the multiplicities of the bonds, the latter depending on the number of surplus bonding electrons available. Molecules with single bonds (H_2, Li_2) have a binding energy which decreases with increasing interatomic distance, since increasing distance reduces the value of the overlap integral. The majority of diatomic molecules H_2, Li_2, N_2, F_2 are diamagnetic, since the occupation of the molecular levels by pairs of electrons leads to compensation of the spin and orbital moments. An interesting exception is the paramagnetic molecule O_2, which possesses an intrinsic magnetic moment equal to two Bohr magnetons. In the oxygen molecule there are altogether 16 electrons, filling successively all the lower molecular levels (Fig. 57) right up to the π^*2p level, which accommodates 2 electrons. This level is doubly degenerate and contains two cells. By a rule analogous to Hund's, these cells are occupied by two electrons with parallel spin orientations. The explanation of the magnetic properties of the oxygen molecule, added to the complete calculation of the energy and the internuclear distance in the hydrogen molecule, was a conspicuous success for molecular quantum theory.

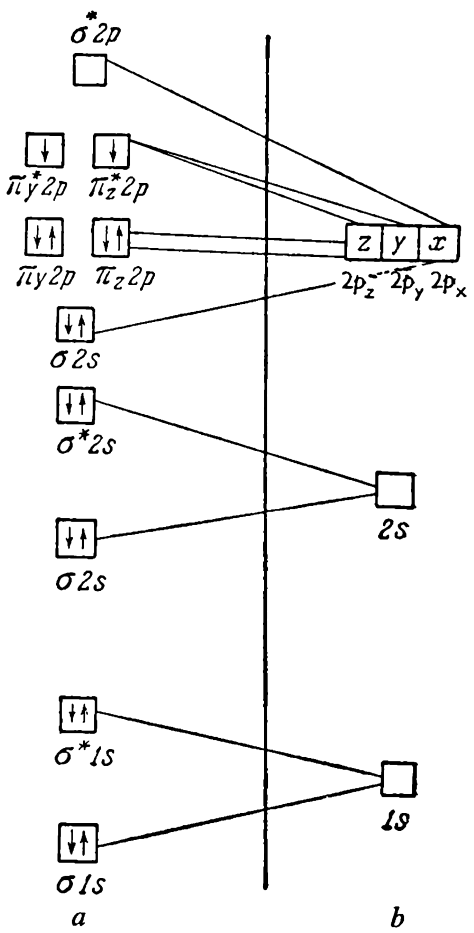


FIG. 57. Sequence of energy levels in diatomic molecules of the type A_2: (*a*) molecular levels, (*b*) atomic levels.

Both electrons on the π^*2p level of the oxygen molecule are antibonding, so the removal of one of them increases the binding energy. Thus the dissociation energy of the molecular ion O_2^+ is greater than that of the neutral molecule O_2. In the nitrogen molecule N_2 the number of valence electrons is 2 less than in the oxygen molecule. All the $2p$-electrons in the N_2 molecule are bonding electrons, so the removal of one of them lowers the dissociation energy of the ion N_2^+ as compared with the molecule N_2.

When multi-electron atoms form molecules there is a difference between the amount of overlapping of the inner and outer electron clouds. Let us consider this in the case of *K*-electrons. When the molecular levels $(\sigma 1s)^2$ $(\sigma^* 1s)^2$ are completely filled the total density of the electrons on these levels is given by the expression

$$2\frac{(\psi_A+\psi_B)^2}{2(1+S)}+2\frac{(\psi_A-\psi_B)^2}{2(1-S)}=\frac{2}{1-S_2}(\psi_A^2+\psi_B^2-2s\psi_A\psi_B). \qquad \ldots [4.10]$$

The denominators of the appropriate terms in [*10*] are determined from the normalisation condition for the wave functions, $N^2 \int \psi^2\, d\tau = 1$. The magnitude of the overlap integral S for the *K*-shells diminishes rapidly with increase in the number of electrons in the atom. In the H_2 molecule $S = 0{\cdot}68$, in the Li_2 molecule $S = 0{\cdot}01$, and for carbon atoms at single bond distance (diamond $R = 1{\cdot}54$ Å) $S = 10^{-5}$. Expression [*10*] transforms into 2 $(\psi_A^2+\psi_B^2)$, when $S = 0$ that is, into the sum of the electron clouds of the invariant *K*-shells of the two atoms. Hence it is possible, and expedient, to divide the electrons into internal (unaffected by a chemical bond) and external (which participate in chemical bonds), the latter being the valence electrons. The scheme of molecular levels, like the scheme of Fig. 56, relates to the valence levels. The energy states of the internal electrons are described by the diagrams of the corresponding atomic levels.

Table 17

Electron Configurations of some Diatomic Molecules and Molecular Ions of Type A_2 in the Ground State

Molecule	$\sigma 1s$	$\sigma^* 1s$	$\sigma 2s$	$\sigma^* 2s$	$\sigma 2p$	$\pi_y 2p$, $\pi_z 2p$	$\pi_z^* 2p$, $\pi_y^* 2p$	$\sigma^* 2p$	Number of valence electrons		Multiplicity of the bond	Dissociation energy eV	Spectroscopic symbol
									Bonding	Anti-bonding			
H_2^+	1								1		½	2·64	$^2\Sigma_g$
H_2	2								2		1	4·72	$^1\Sigma_g$
He_2^+	2	1							2	1	½	(2·5)	$^2\Sigma_u$
He_2	2	2							2	2	0	0	$^1\Sigma_g$
Li_2	2	2	2						2	0	1	1·4	$^1\Sigma_g$
Be_2	2	2	2	2					2	2	0		$^1\Sigma_g$
N_2^+	2	2	2	2	1	2, 2			7	2	2½	6·35	$^2\Sigma_g$
N_2	2	2	2	2	2	2, 2			8	2	3	7·38	$^1\Sigma_g$
O_2^+	2	2	2	2	2	2, 2	1		8	3	2½	6·48	$^2\Pi_g$
O_2	2	2	2	2	2	2, 2	1, 1		8	4	2	5·08	$^3\Sigma_g$
F_2^+	2	2	2	2	2	2, 2	2, 1		8	5	1½		$^2\Pi_g$
F_2	2	2	2	2	2	2, 2	2, 2		8	6	1	(2·8)	$^1\Sigma_g$
Ne_2	2	2	2	2	2	2, 2	2, 2	2	8	8	0		$^1\Sigma_g$

The invariance of the states of the inner electrons belonging to atoms in molecules and crystals confirms the correctness of our initial approximation, that is of the physical model which forms the foundation of the method of linear combination of atomic orbitals.

Binding Energies and Interatomic Distances in Diatomic Homopolar Molecules

Table 18 gives the values of the binding energies and interatomic distances in diatomic molecules formed by elements of the B-subgroups, and also of the IA-subgroup.

Table 18

Dissociation Energies *D* and Interatomic Distances *r* in Homopolar Diatomic Molecules of Elements of the B-subgroups

	IA	IIB	IIIB	IVB	VB	VIB	VIIB
	Li	Be	B	C	N	O	F
D, eV	1·14		3·0	3·6	7·38	5·08	2·8
r, Å	2·67		2·59	1·31	1·09	1·21	1·45
	Na	Mg	Al	Si	P	S	Cl
D, eV	0·76				5·03	3·6	2·48
r, Å	3·08				1·89	1·90	2·01
	K	Zn	Ga	Ge	As	Se	Br
D, eV	0·51				3·96	2·7	1·97
r, Å	3·92					2·19	2·28
	Rb	Cd	In	Sn	Sb	Te	I
D, eV	0·49	0·09			(3·7)	2·3	1·54
r, Å						2·59	2·66
	Cs	Hg	Te	Pb	Bi		
D, eV	0·45	0·08		(0·7)	3·34		
r, Å							

The interatomic distances in the molecules increase as we go through the members of each subgroup in order of increasing atomic number, and at the same time the dissociation energies correspondingly diminish because of the decrease in the overlap integral. As the principal quantum number increases, the energy levels for electron states come closer together and the corresponding excitation energies become less. This allows atoms of the heavier elements of the B-subgroup to use higher valencies in addition to their fundamental valencies. The excitation of a new valency increases the original valency by an even number. Thus halogens are known to have, in addition to their basic valency 1, higher valencies 3, 5 and 7, as shown by the existence of the molecules ClF_3, IF_3, Cl_2O_7. The elements of the VI B-subgroup—sulphur, selenium, tellurium—exhibit valencies 4 and 6 as well as their basic valency 2. The elements of the V B-subgroup—arsenic, antimony, bismuth—have a valency 5 as well as their basic valency 3.

In the series of molecules F_2, O_2, N_2, the dissociation energies increase since the multiplicity of the bond increases: $F-F$, $O=O$, $N\equiv N$. The trivalent nitrogen atom is more active chemically than the divalent oxygen atom. However molecular nitrogen is less active than molecular oxygen, and the latter is less active than molecular fluorine. The explanation of this is the difference between the multiplicities of the bonds and the corresponding difference in the dissociation energies of the molecules in question. In the C_2 molecule there is, in contrast to N_2, not an increase but a decrease in the multiplicity of the bond. The bonds in the C_2 molecule are double, not quadruple, and correspond to the divalent state of the carbon atom although the carbon atom is basically tetravalent. In the molecule C_2 (as also in the molecule of carbon monoxide CO) this valency is not found. The absence of molecules of the type $C \equiv C$ shows that it is impossible to form quadruple bonds; this is a consequence of the space quantisation rule. The greatest possible multiplicity of a chemical bond is three.

The existence of diatomic molecules of the alkali metals in the gaseous phase has been established spectroscopically. The dissociation energies and internuclear distances in these molecules vary with their positions in the subgroup, in order of increasing atomic number, in the same way as in the molecules of halogens. In both cases the bonds in the molecules are covalent two-electron (single) bonds. The difference in structure between the valence shells of halogen and alkali metal elements leads to a fundamental difference in the saturability of the bonds. This is shown by the different structure of the crystals formed by halogens and alkali metals.

Polyatomic Molecules

In polyatomic molecules the valence electrons move in orbitals which encompass several nuclei. The molecular orbitals in this case represent a linear combination of atomic orbitals $\psi_A, \psi_B, \psi_C \ldots$,

$$\psi = c_A\psi_A + c_B\psi_B + c_C\psi_C + \ldots$$

We now take into account the condition for the formation of the most effective bonds (maximum overlapping of the atomic wave functions). This condition means that not all sets of atomic functions are equally important, the most important being linear combinations of the atomic functions of nearest neighbour atoms. We can illustrate this by the example of the water molecule. Hydrogen atoms have a single valence 1*s*-electron each. Oxygen atoms derive their valency in the first instance from two electrons ($2p_x$ and $2p_y$), which produce the basic divalent state of oxygen. The electron clouds (Fig. 58) of these electrons are mutually perpendicular. The greatest overlapping of the wave functions of the oxygen atom and the hydrogen atoms occurs when the hydrogen atoms are orientated in the directions of the *X*- and *Y*-axes associated with the oxygen atom. This brings about the formation of the triatomic molecule H_2O, which possesses a triangular structure.

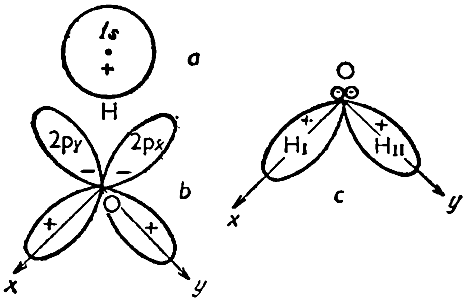


FIG. 58. Structure of the water molecule.

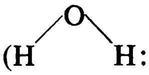
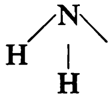
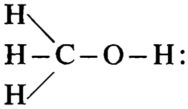
Since oxygen atoms are more electronegative than hydrogen atoms, the O–H bonds have a dipole moment with the negative end of the dipole on the oxygen atom. Vector addition of the dipole moments of the O–H bonds gives the water molecule a significant electric moment; this explains its polar properties. The repulsion of the positive ends of the dipoles (protons) increases the valence angle from 90° to 104·5°.† In the water molecule the oxygen atom occupies a dominant position and is linked with two hydrogen atoms. The length of the O–H bond is 1·07 Å, which is roughly equal to the sum of the covalent radii of the oxygen and hydrogen atoms:

---

† The increase in the valence angle may be explained as a result of the sharing of the *s*-state in the bonds and the addition of tetrahedral sp^3-hybridisation (see below).

$$r_O + r_H = 0{\cdot}66 + 0{\cdot}37 = 1{\cdot}03 \text{ Å}.$$

The distance between the protons in the water molecule ($\sim$1·6 Å) considerably exceeds the corresponding distance in the hydrogen molecule H–H (0·74 Å). This shows that in the water molecule there is no direct interaction between the protons comparable in magnitude to the interaction between the hydrogen and oxygen atoms. The basic rôle in the water molecule is played by the O–H bonds and the valence state of the oxygen atom (p^2-electrons) determines the triangular structure of the water molecule.

The account we have given shows that in polyatomic molecules it is possible to divide the atoms into those bound directly together and those not so bound, and to consider the molecule as a system formed by an arrangement of individual bonds. This corresponds to the representation of the structure of molecules by means of valence lines (: water molecule, : ammonia molecule, : methyl alcohol molecule etc.). This kind of description of the structure is simplified, but in a number of cases quite satisfactory, in so far as the properties of individual pairs of bound atoms in the same valence state of these atoms (properties of bonds) remain almost identical in different molecules. The properties which characterise the bonds are: structural characteristics (interatomic distances, coordinations, valence angles), bond strength (energy, frequencies of oscillation), and the electric, magnetic, optical and other properties of the molecules. However, the analysis of molecules into additive arrays of individual bonds is not possible in all cases; for example, it is impossible in molecules with multiple coupled bonds (benzene), and also in crystals with electronic conductivity (graphite, metals), where the properties are determined by the whole set of valence electrons.

Let us consider first the structure of molecules in which such an analysis is possible, for example, the water molecule. In this molecule, in place of a combination of four atomic orbitals,

$$c_1\mathrm{H}_1\,(1s) + c_2\mathrm{H}_2\,(1s) + c_3\mathrm{O}\,(2p_x) + c_4\mathrm{O}\,(2p_y),$$

we can take two identical combinations of two atomic orbitals,

$$\psi_{\mathrm{I}} = H_1\,(1s) + \lambda\mathrm{O}\,(2p_x) \quad \text{and} \quad \psi_{\mathrm{II}} = \mathrm{H}_2\,(1s) + \lambda\mathrm{O}\,(2p_y), \qquad \ldots [4.11]$$

corresponding to the two O–H bonds. We note that in consequence of the orthogonality of the p_x and p_y functions the overlap integral

$$\int \psi(p_x)\,\psi(p_y)\,d\tau = 0.$$

In elements of the VB-subgroup the basic valence state is determined by three p-electrons, which form 3 mutually perpendicular bonds. So the ammonia molecule has a pyramidal structure. Owing to the repulsion of the protons the valence angle is slightly enlarged, as in the water molecule.

Hybridisation of the Electron States (Directed Valencies)

We have considered bonds formed by atoms possessing several valence electrons of the same type (p^2 and p^3). In cases when the atoms combined are of the same type, it is natural to expect all the bonds to be equivalent.

Let us now pass on to consider bonds formed by atoms which have valence electrons of different types *s*, *p*, *d*. We shall begin with the case when there are only *s*- and *p*-electrons in the valence shell. Let the valence shell contain one *s*- and one *p*-electron (the valence *sp*-state, for example the excited atoms Be*, Mg*, Zn* etc.); to be specific, we shall take the *p*-electron to be a p_z-electron. The *s*- and *p*-electrons can form two bonds, as is confirmed by the existence of the molecules $ZnCl_2$, $ZnBr_2$ etc., but the difference between the valence electrons points to the possibility that the bonds formed by them may not be equivalent. In fact the *s*-electron forms a weaker bond than does the *p*-electron. If, starting from the condition of maximum overlapping, we measure the bond strength by the maximum value of the angular part of the wave function, and take the strength of the *s*-bond as unity, then the strength of the *p*-bond will be $\sqrt{3} = 1{\cdot}732$, which is considerably greater than that of the *s*-bond. On the basis of the principle of superposition we shall try to form from the initial *s* and p_z wave functions two new wave functions

$$\psi_1 = a_1 s + b_1 p_z; \quad \psi_2 = a_2 s + b_2 p_z, \qquad \ldots [4.12]$$

where a_i, b_i are coefficients determined by (1) the normalisation condition for the wave functions, $\int \psi_i^2 d\tau = 1$, (2) the orthogonality condition $\int \psi_i \psi_k d\tau = 0$, and (3) the equivalence condition for the functions ψ_i and ψ_k. We shall call wave functions formed from functions of different types and satisfying the equivalence condition *hybrid* electron functions. From the sphericity of the *s*-state and the equivalence condition it follows that $a^2 = \frac{1}{2}$:

$$a_1 = a_2 = a = \frac{1}{\sqrt{2}}. \qquad \ldots [4.13]$$

From the normalisation and orthogonality conditions for the *s*- and *p*-functions we have

$$a^2 + b_1^2 = 1, \quad \text{so} \quad b_1 = \frac{1}{\sqrt{2}}. \qquad \ldots [4.14]$$

From the orthogonality condition for the hybrid functions ψ_1 and ψ_2 we have

$$\int \psi_1 \psi_2 d\tau = a^2 + b_1 b_2 = \frac{1}{2} + \frac{b_2}{\sqrt{2}}, \text{ so } b_2 = -\frac{1}{\sqrt{2}}. \qquad \ldots [4.15]$$

Substituting into [*12*] the values of the coefficients from [*13*], [*14*] and [*15*], we obtain the expression for the hybrid functions:

$$\left.\begin{aligned} \psi_1 &= \frac{1}{\sqrt{2}} s + \frac{1}{\sqrt{2}} p_z = \frac{1}{\sqrt{2}} + \sqrt{\frac{3}{2}} \cos\theta, \\ \psi_2 &= \frac{1}{\sqrt{2}} s - \frac{1}{\sqrt{2}} p_z = \frac{1}{\sqrt{2}} - \sqrt{\frac{3}{2}} \cos\theta. \end{aligned}\right\} \qquad \ldots [4.16]$$

It is not difficult to see that both the hybrid wave functions give the same electron density distribution: one in the direction of the positive Z-axis, the other in the direction of the negative Z-axis (the valence angle is 180°). This explains the linear form of the molecule of type $ZnCl_2$ and the equivalence of the two bonds $Zn-Cl_I$ and $Zn-Cl_{II}$. The bond strength connected with the hybrid functions,

$$(\psi_1)_{\max} = (\psi_2)_{\max} = \frac{1+\sqrt{3}}{\sqrt{2}} = 1{\cdot}932, \qquad \dots [4.17]$$

considerably exceeds the strengths of the bonds formed by either s-, or p-electrons. This shows how the potential energy of hybrid bonds compares in effectiveness with that of bonds formed by electrons in the invariant s- and p-states.

We can now find hybrid functions for sp^2-electrons. Suppose that they are p_x- and p_y-electrons, so that

$$\left.\begin{aligned} \psi_1 &= a_1 s + b_1 p_x + c_1 p_y \\ \psi_2 &= a_2 s + b_2 p_x + c_2 p_y \\ \psi_3 &= a_3 s + b_3 p_x + c_3 p_y. \end{aligned}\right\} \qquad \dots [4.18]$$

From the equivalence condition $a_i^2 = \frac{1}{3}$, $\quad a_1 = a_2 = a_3 = a = \frac{1}{\sqrt{3}}$.

Let the function ψ_1 have a maximum in the direction of the X-axis; then the function p_y will play hardly any part in it, so we can put $c_1 = 0$. From the normalisation condition $b_1 = \sqrt{\frac{2}{3}}$. From the orthogonality condition $\int \psi_1 \psi_2 d\tau = 0$ we obtain $b_2 = -\frac{1}{\sqrt{6}}$. The normalisation condition applied to ψ_1 gives $c_2^2 = 1 - a_2^2 - b_2^2 = \frac{1}{2}$.

Similarly we find $b_3 = b_2$ and $c_3^2 = c_2^2$, whence $c_2 = +\frac{1}{\sqrt{2}}$ and $c_3 = -\frac{1}{\sqrt{2}}$ (or vice versa). Introducing the values found for the coefficients into [*18*], we obtain

$$\left.\begin{aligned} \psi_1 &= \frac{1}{\sqrt{3}} s + \sqrt{\frac{2}{3}} p_x = \frac{1}{\sqrt{3}} + \sqrt{2} \cos \phi, \\ \psi_2 &= \frac{1}{\sqrt{3}} s - \frac{1}{\sqrt{6}} p_x + \frac{1}{\sqrt{2}} p_y = \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}} \cos \phi + \sqrt{\frac{3}{2}} \sin \phi, \\ \psi_3 &= \frac{1}{\sqrt{3}} s - \frac{1}{\sqrt{6}} p_x - \frac{1}{\sqrt{2}} p_y = \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{2}} \cos \phi - \sqrt{\frac{3}{2}} \sin \phi. \end{aligned}\right\} \qquad \dots [4.19]$$

It is not difficult to see that the function ψ_2 has its maximum value in the direction for which $\phi = 120°$, and the function ψ_3 in the direction with $\phi = 240°$. The three hybrid sp^2-bonds are uniformly distributed in a plane, form valence angles of 120° with one another and are orientated towards the vertices of an equilateral triangle at the centre of which the atom under consideration is situated. Such is the structure of the plane triangular molecules BF_3, BCl_3 etc. The strength of all these bonds is the same and is

$$(\psi_1)_{\max} = (\psi_2)_{\max} = (\psi_3)_{\max} = \frac{1}{\sqrt{3}} + \sqrt{2} = 1{\cdot}991, \qquad \dots [4.20]$$

which exceeds the strengths not only of the *s*- and *p*-states, but also of the hybrid *sp*-functions.

In the *s*- and *p*-cells there may be up to four valence electrons, forming an sp^3-group. The hybrid functions for them will be of the type

$$\psi_i = a_i s + b_i p_x + c_i p_y + d_i p_z. \quad (i = 1, 2, 3, 4). \qquad \ldots [4.21]$$

From the equivalence condition $a_i^2 = \frac{1}{4}$:

$$a_1 = a_2 = a_3 = a_4 = a = \tfrac{1}{2}. \qquad \ldots [4.22]$$

The normalisation conditions give 4 relationships of the type

$$\tfrac{1}{4} + b_i^2 + c_i^2 + d_i^2 = 0. \qquad \ldots [4.23]$$

The orthogonality conditions give 6 relationships of the type

$$\tfrac{1}{4} + b_i b_k + c_i c_k + d_i d_k = 0. \qquad \ldots [4.24]$$

Let us suppose that the function ψ_1 has a maximum in the direction of the *Z*-axis; then the functions p_x and p_y hardly participate in the formation of this function. This gives two additional relationships:

$$b_1 = c_1 = 0. \qquad \ldots [4.25]$$

These 16 conditions are sufficient to determine all the coefficients in [*21*]. Thus, from the normalisation condition for ψ_1 we find $b_1 = \sqrt{1-a^2} = \frac{\sqrt{3}}{2}$, consequently

$$\psi_1 = \tfrac{1}{2}s + \frac{\sqrt{3}}{2}\,p_z = \frac{1}{2} + \frac{3}{2}\cos\theta. \qquad \ldots [4.26]$$

Analogous expressions may be obtained also for the remaining hybrid functions. Analysis of the positions of the maxima for the hybrid functions shows that they are uniformly distributed in space and orientated along the diagonals of a regular tetrahedron. The valence angles between the bonds are 109° 28′. Such is the structure of the methane molecule CH_4. The strength of the hybrid sp^3-bonds is 2, as can be seen from [*26*], and is greater than the strengths of all the bonds previously considered.

We can now show that among all the possible bonds that can be formed by *s*- and *p*-electrons, the hybrid sp^3-bonds have the maximum strength. We write the expression for the hybrid function [*26*] (in the direction of the *Z*-axis) in the form

$$\psi_1 = a_1 s + d_1 p_z = a_1 + \sqrt{3(1-a_1^2)}. \qquad \ldots [4.27]$$

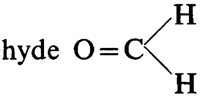
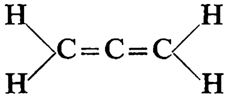
We find the value of the coefficient a_1 from the condition for a maximum $\frac{d\psi_1}{da} = 0$, which leads to the equation

$$1 - \frac{\sqrt{3a_1}}{(1-a_1^2)^{\frac{1}{2}}} = 0, \qquad \ldots [4.28]$$

from which we obtain $a_1 = \frac{1}{2}$. This coincides with the value for a found previously from the equivalence condition.

If the tetrahedral hybrid functions are written not in triangular form as previously (the Z-axis was oriented along a threefold axis of symmetry of the tetrahedron), but in the form of components along cubic axes (the X, Y, Z axes are orientated along the twofold axes of symmetry of the tetrahedron), then the expressions for the hybrid functions will take on a more symmetrical form:

$$\left.\begin{aligned} \psi_1 &= \psi_{[111]} = \tfrac{1}{2}(s+p_x+p_y+p_z), \\ \psi_2 &= \psi_{[1\bar{1}\bar{1}]} = \tfrac{1}{2}(s+p_x-p_y-p_z), \\ \psi_3 &= \psi_{[\bar{1}1\bar{1}]} = \tfrac{1}{2}(s-p_x+p_y-p_z), \\ \psi_4 &= \psi_{[\bar{1}\bar{1}1]} = \tfrac{1}{2}(s-p_x-p_y+p_z). \end{aligned}\right\} \quad \ldots [4.29]$$

The hybrid functions thus found, which are formed from the s- and p-wave functions, describe the basic valence states of atoms with s- and p-electrons (atoms of elements belonging to the B-subgroups). Let us consider this in the case of carbon atoms, the majority of whose compounds are characterised by three valence states, described by hybrid sp^3-, sp^2- and sp-functions (σ-bonds). sp^3-hybridisation occurs in compounds involving the tetrahedral carbon atom, which was discovered in organic chemistry in the second half of the last century (aliphatic compounds). Tetrahedral bonds are a feature of one of the modifications of carbon—the diamond crystal—and also of silicon, germanium and grey tin crystals. In these crystals the atoms form four σ-bonds; this determines the coordination number in the diamond type structure, $k = 4$. sp^2-hybridisation takes place in molecules of the formaldehyde  $O{=}C\begin{smallmatrix}H\\H\end{smallmatrix}$ type, and in molecules of aromatic compounds (benzene C_6H_6 etc.). In these compounds the carbon atom forms three coplanar σ-bonds; this determines the coordination number of the carbon atom, $k = 3$. The fourth valence p_z-electron of the carbon atom forms an additional π-bond. sp^2-hybridisation is characteristic of another modification of carbon—graphite. sp-hybridisation occurs in molecules of carbon dioxide $O{=}C{=}O$ and carbon disulphide CS_2, in the propadiene molecule  $\begin{smallmatrix}H\\H\end{smallmatrix}C{=}C{=}C\begin{smallmatrix}H\\H\end{smallmatrix}$ and in other molecules in which the carbon atom forms two double bonds. In these compounds the carbon atom forms two linear σ-bonds; this agrees with the linear form of the corresponding molecules as observed experimentally, and determines the coordination number of the carbon atom, $k = 2$. The two other valence p-electrons of the carbon atom form two additional π-bonds.

The explanation of the rôle of hybrid states in the formation of interatomic bonds (Pauling, 1928) was a conspicuous success for the theory and enabled us to understand the reason for directed bonds. The hybrid states considered above are in essence different forms of spatial distribution of the valence electron density, compatible with the laws of motion of electrons in atoms (conservation of angular momentum, normalisation of the wave functions, orthogonality). During interatomic interaction the electron density of the atom is redistributed in such a way that the gain in interaction energy is a maximum.

The theory of hybrid states has made it possible not only to explain the previously known valence states, as in the case of carbon compounds, but also to predict new types of coordination, which have since been discovered experimentally. A further question concerns the valence states of atoms containing d-electrons as well as s- and p-electrons, that is of the atoms of transition metals. Consideration of the angular parts of the *d*-wave functions shows that the maximum strength of the bond formed by the *d*-electron, $\sqrt{5} = 2{\cdot}236$, exceeds the bond strengths of the *s*- and *p*-electrons and their hybrid functions. The maximum strength of the hybrid function which can be formed from *s*-, p- and *d*-functions is 3. Participation of *d*-electrons in hybridisation thus raises the bond strength.

Out of the large number of hybrid functions formed from *s*-, *p*- and *d*-functions, we shall indicate some of the most important. d^3s-hybridisation leads to tetrahedral bonds and is similar to sp^3-hybridisation. dsp^2-hybridisation leads to four coplanar bonds, forming valence angles of 90° and oriented along the diagonals of a square; the strength is 2·694. A similar square configuration was found in the complex ions of transition metals of the type $[PtCl_4]^{-2}$, $[Ni(CN)_4]^{-2}$ etc. d^2sp^3-hybridisation leads to six mutually orthogonal bonds oriented along the diagonals of an octahedron, the strength being 2·923. Octahedral coordination is very common in the complex ions $[PtCl_6]^{-2}$, $[Fe(CN)_6]^{-3}$ etc.

The theory of hybrid states reveals the possible types of configuration but does not allow us to make a definite choice between them, since the bond strength is a very poor indication of the binding energy.

Free Electron Pairs

Not all valence electrons in a given atom take part in the formation of a bond. Apart from the bonding and antibonding electrons there may be unused electrons in the form of electron pairs with compensated spins. These free electron pairs may take part in the formation of additional bonds and are essential to an understanding of the properties of those molecules whose atoms contain them.

We shall illustrate the existence of electron pairs in the case of the molecules HCl, H_2O and NH_3. The HCl molecule contains a σ-bond formed by the 1*s*-electron of the H atom and the $3p_x$-electron of the Cl atom. In the valence shell of the Cl atom there are in addition electron pairs $3p_y$ and $3p_z$, which could form two π-bonds. In the hydrogen atom a π-bond could be formed on account of the vacant 2*p*-level, but its energy is so high that there is in fact no possibility of forming π-bonds in the HCl molecule. Thus in this molecule, taking into account the presence of three s^2-electrons, the Cl atom has three free electron pairs. Similarly the oxygen atom in the water molecule has two free electron pairs, and the nitrogen atom in the ammonia molecule has one.

2. The Metallic Bond

The Formation of a Crystal as a Consequence of an increase in the Degree of overlapping of Atomic Wave Functions

Overlapping of wave functions takes place not only when two atoms approach each other but also when a greater number of atoms come together. Fig. 59 shows the overlapping of the 2*s*-wave functions of lithium atoms separated by distances equal to the interatomic distances in the lithium crystal. The energy gain after an increase in the number of overlaps produces a tendency for each of the atoms to form the largest possible number of bonds, that is, to surround itself with the greatest possible number of neighbours. This must lead to the formation of structures extending infinitely throughout space with a dense packing of particles, since such structures have the highest possible coordination numbers. As experiment shows, the majority of metals actually crystallise in several simple, highly symmetrical structures with large coordination numbers (Ch. 6). Whether the atoms form an unbounded crystal or separate molecules (polyatomic, diatomic or monatomic) depends on which of these states corresponds to the energy minimum of the system.

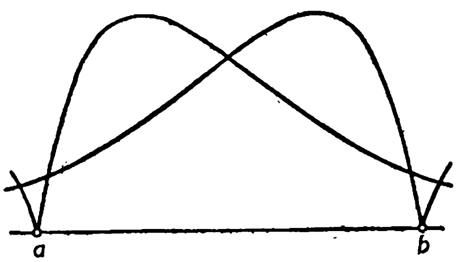


FIG. 59. Overlapping of wave functions in the lithium crystal.

Level Scheme of Crystals. Energy Bands

We shall represent the formation of a crystal as the successive addition of atoms to an arbitrarily chosen first atom. Consider a particular energy level: in a system of two atoms the original level splits into two levels, in the case of three into three, etc. In a system of N atoms the initial N-fold degenerate level splits into N levels (Fig. 60). Since the total number of atoms in a crystal is high (in 1 cc of a metal there are about 10^{22}-10^{23} atoms), neighbouring levels are very close together and form a quasi-continuous zone or energy band. In contrast to the line spectrum of individual atoms, the energy spectrum of the upper (valence) filled levels of a crystal consists of separate bands in which the energy levels are successively filled up in accordance with the Pauli principle.

The formation of quasi-continuous bands in the energy spectrum of a crystal explains the appearance of continuous emission spectra in crystalline substances, and also in amorphous bodies and liquids. Immediate confirmation of the zone structure of the crystalline energy spectrum is provided by X-ray emission and absorption spectra. In Fig. 61 AB is one of the lower energy levels of atoms in the crystal, which in practice hardly varies. $CDGF$ is a valence band of the crystal in which all the levels are filled from the bottom of the band DG to the uppermost filled edge PP. X-ray emission spectra arise during transitions from one of the filled levels to a lower free level. Because of the finite depth, ΔE, of the filled zone, the spectral lines possess a finite breadth and definite shape. The narrow line of the X-ray emission spectrum for a gas (Fig. 62*a*) changes into a broad line for a crystal (Fig. 62*b*). The intensity distribution in terms of the breadth of the lines is

$$I(\nu) = I(E) \approx p(E) \cdot n(E) \qquad \ldots [4.30]$$

and is given by the product of two factors: $n(E)$—the density of energy levels in the band, and $p(E)$—the probability that an electron will make the transition from one of the occupied levels to a lower empty level. An X-ray *absorption* spectrum arises during the transfer of an electron from one of the lower levels to an *unfilled* part of the band. On the continuous absorption curve which varies, as shown by theory, according to an approximately (arctan ν) law there appears an 'absorption edge'

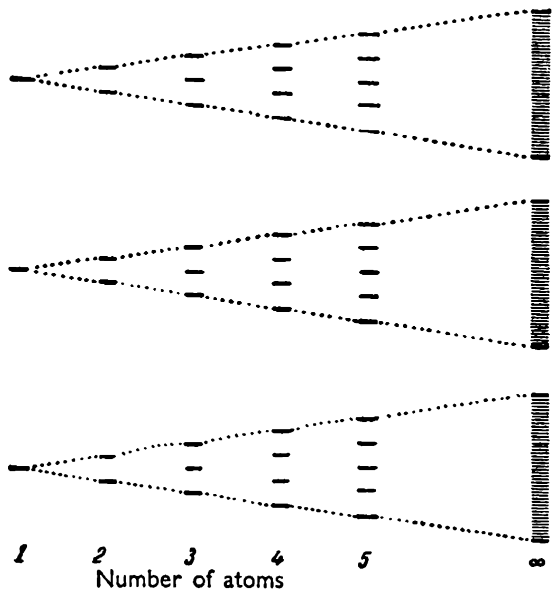


FIG. 60. Removal of degeneracy in a system of coupled atoms.

(Fig. 62*c*), the position and shape of which are determined by the energy level density and the probability of transitions to the unfilled part of the zone, $I'(E) = p'(E) \cdot n'(E)$. The curve $I'(E)$ is shown dotted in Fig. 62*c*. The study of the X-ray spectra of crystals enables us to obtain important energy data about them.

The zone structure of the energy spectrum in a crystal arises from the solution of the idealised quantum mechanical problem of the motion of an electron in a periodic field (Ch. 1, §8).

Features of Metallic Binding

Metallic atoms are those which form structures in which the number of bonds, determined by the number of closest neighbours, exceeds the number of valence electrons ($k > v$). Thus, in crystals of the alkali metals, the atoms of which are monovalent, each atom forms $8+6 = 14$ bonds. The 8 bonds with atoms of the first coordination sphere are stronger than the 6 bonds with atoms of the second coordination sphere, the radius of which is only 14% more than the radius of the first sphere. The mean number of bonds can be taken as ~ 12 (which corresponds to the coordination number in close packing of spheres). On time average an atom gives $\frac{1}{12}$ of an electron to each bond, and two neighbouring atoms in conjunction give $\frac{1}{6}$ of an electron, instead of the two electrons which enter into a 'normal' identical covalent

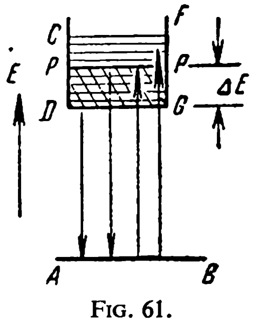


FIG. 61.
Origin of X-ray spectra.

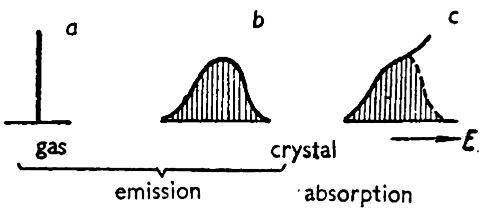


FIG. 62.
X-ray emission and absorption spectra.

bond. In the structures of metals we are dealing with fractional bonds, with multiplicity less than unity. A similar situation obtains also in the molecular hydrogen ion H_2^+, considered previously, for which the multiplicity of the bond is $\frac{1}{2}$ ('single-electron bond').

The molar sublimation energies of the alkali metals are 1·5-2 times the dissociation energies of their diatomic molecules.

Sublimation Energies Q and Shortest Inter-Atomic Spacings r_0 in crystals of the alkali metals

	Li	Na	K	Rb	Cs
Q, eV	1·7	1·1	0·85	0·81	0·80
r_0, Å	3·03	3·72	4·50	4·95	5·3

However the energy of each bond in a metallic crystal is less than in a molecule. In the lithium crystal the energy of a single bond amounts on average to $\frac{1}{12} \times 1{\cdot}7 \approx 0{\cdot}14$ eV as against 1·14 eV in the Li_2 molecule. The weakening of the bonds in metallic crystals is confirmed by the increase in the interatomic spacings, which are 10-15% greater than those in diatomic molecules. The weakening of each bond in metals compared to the single bond is more than compensated by the increase in the number of bonds, so the formation of a coordination crystal is energetically more favourable than the formation of the corresponding set of molecules with low coordination numbers.

The metallic bond is unsaturated. In such a bond there is usually a tendency to form structures with the highest possible coordination numbers such as are found in dense packing of spheres. However it cannot be said that the field of a metallic atom has complete spherical symmetry. This is indicated by the fact that crystals of cubic metals do not obey Cauchy's relations, which equate certain elastic constants in crystals where central forces act between the particles. It is also indicated by the fact that metals crystallise not only in the closest packing (packing of spheres), but also in other close packings (the body-centred cubic structure).

The fundamental and characteristic physical properties of metals—high electrical conductivity and metallic lustre—confirm the theory that the valence electrons are collectivised and can move throughout the volume of the whole crystal. These properties had been considered evidence for the existence of free electrons in metals long before Drude (1905), in the very earliest electron theory of metals, ascribed to them the properties of an ideal gas. However this turned out to be incorrect, and in the later development of the theory of metals the classical Boltzmann statistics had to be replaced by quantum statistics taking into account the Pauli principle.

Metallic Hydrogen

The symmetry of the wave function for the hydrogen atom is the same as for the lithium atom. Starting from the concept of maximum overlapping of wave functions, we would find it interesting to examine the possibility of forming hydrogen with a high coordination structure like the structure of the alkali metals—of forming, that is, hydrogen in the metallic state. The calculated energy curve for such a hypothetical substance (Wigner and Huntingdon, 1935) has a minimum, the depth of which, with a correction for the zero point energy, gives the crystal energy as 10·6 kcal./mole and the density as 0·59 g./cm.^{3}. The equilibrium state corresponds to a value $a = 1{\cdot}78$ Å for the length of the edge of the unit cube and to an interatomic distance of $r = 1{\cdot}54$ Å.

Metallic hydrogen can exist in principle, but only by use of higher quantum states. So at normal pressure (and low temperature) the crystal of molecular hydrogen H_2 is more stable. The energy of the molecular modification amounts to 52·4 kcal./mole at a density of 0·087 g./cm.^{3}. With increasing pressure the energy of the molecular modification increases more rapidly than that of the metallic modification. An approximate calculation shows that the point of transition from one modification to the other lies at a pressure of the order of 300,000 atm.

The possibility of hydrogen existing in the metallic state is indicated, as Ubbelohde observed (1937), by the hydrides of certain transition metals which conserve their metallic properties and have the same lattice as the original metal. Palladium has a great capacity for absorbing hydrogen. The formation of the hydride of palladium, which contains 39 at. % of hydrogen, is accompanied by an increase of 3·5% in the lattice constant of palladium. If we suppose that the coefficient of compressibility of palladium hydride is roughly the same as for metallic palladium, we can calculate that this corresponds to an increase of $2{\cdot}75 \times 10^5$ atm. in the internal pressure, in agreement with the results of theoretical calculations.

3. The Heteropolar Bond

The Quantum Treatment of Heteropolar Compounds

Heteropolar molecules and crystals are formed from atoms with different electronegativities and are a very common class of compounds. For heteropolar molecules AB $|\lambda| \neq 1$ in formula [*3.22*]. However the conditions for the formation of effective bonds (approximate energy equality, maximum overlap and identical symmetry about the bond line of the initial atomic orbitals) remain the same and enable us to choose the most important orbitals ψ_a and ψ_b. The division into σ- and π-bonds also retains its significance but because there is no centre of symmetry in the molecule, the division of its levels into levels of types *u* and *g* does not take place. The molecular levels for divided atoms transform into non-coincident atomic levels in view of the difference between the atoms *a* and *b*.

The most fundamental difference between heteropolar and homopolar molecules lies in the appearance of electric dipole moments. If $|\lambda| > 1$, then in the molecular orbital $\psi_a + \lambda\psi_b$ the atomic orbital ψ_b has greater weight than the orbital ψ_a. The valence electron moves in such a way that for most of the time it is near atom B. Thus, when a molecule is formed, the electron density is redistributed such that it is increased near atom B and reduced near atom A. This produces in the molecule a dipole moment determined by the magnitude of λ. To a first approximation we can take it that the electron densities corresponding to the functions ψ_A and ψ_B enter in the proportion $1:\lambda^2$. If the centres of gravity of the electron charges $\frac{1}{1+\lambda^2}$ and $\frac{\lambda^2}{1+\lambda^2}$ bound to the atoms A and B, coincide with the positions of the nuclei A and B, then the molecule's electric dipole moment μ, created by the redistribution of the valence electrons around the nuclei, is given approximately by the formula

$$\mu = \left(\frac{\lambda^2-1}{\lambda^2+1}\right) eR. \qquad \ldots [4.31]$$

A more accurate expression for μ can be obtained by using the function ψ to calculate the electron charge distribution more accurately.

Dipole Moments of Heteropolar Molecules and Electronegativities of the Atoms

From formula [*31*], knowing the factor λ which determines the polarity of the bond, we can find the dipole moment of the molecule. Since it is difficult to calculate λ accurately, the opposite procedure is often followed and λ found from the experimental value for μ. Table 19 gives the dipole moments of certain polar molecules and the values of λ calculated from them, the quoted values referring to the halogen atoms, which contain the negative end of the dipole. The fourth column gives the effective charges of the dipoles, these charges being concentrated on the nuclei and expressed in fractions of the electronic charge. For homopolar molecules with purely covalent bonds these charges are zero, while for the limiting case of ionic monovalent molecules these charges must be equal to $\pm 1{\cdot}0e$.

The bond in heteropolar molecules is sometimes considered as an intermediate ionic-covalent bond and the ionic bond percentage is measured by the effective charges of the dipoles. In the hydrogen halide molecules the ionic character of the bond increases considerably as we pass from HI to HF. The polarity of the bond depends on the difference between the electronegativities of the atoms. The dipole moments of the molecules are comparable to the differences of the electronegativities: $\mu \sim x_A - x_B$, the latter being calculated according to the data of Table 14 (p. 84). For a number of hydrogen halides there is a good agreement between the figures (Table 19).

Table 19

Dipole Moments of Certain Molecules

Molecule	$\mu_{exp} \times 10^{18}$ e.s.u.	R, Å	$\frac{\mu}{eR}$	λ calc	$x_A - x_B$
KCl	6·3	2·79	0·47	2·0	2·2
HF	1·91	0·92	0·43	1·9	1·9
HCl	1·03	1·27	0·17	1·3	0·9
HBr	0·78	1·41	0·11	1·2	0·7
HI	0·38	1·61	0·05	1·06	0·4

Approximate Calculation of the Energy of the Ionic Bond

Typical ionic compounds are the alkali halides. We shall find the energy required to form a molecule from the neutral atoms.

The first stage consists in the formation of ions. The formation of the positive ion Na^+ requires the expenditure of energy equal to the ionisation potential, I_{Na}, of the sodium atom. The formation of the negative ion Cl^- is accompanied by a gain in energy, E_{Cl}, corresponding to the energy of affinity of the electron to the chlorine atom. In all, the process of ion formation requires the expenditure of an amount of energy equal to

$$-I_{Na} + E_{Cl} = -5{\cdot}1 + 3{\cdot}7 = -1{\cdot}4 \text{ eV}$$

The second stage consists in electrostatic attraction between the ions and is accompanied by a considerable energy gain. The potential created by a univalent ion at a distance $r = 2{\cdot}5$ Å, equal to the interatomic distance, in the molecule NaCl will (according to the spectroscopic data) be

$$V = \frac{e}{r} = \frac{4{\cdot}8\times 10^{-10}}{2{\cdot}5\times 10^{-8}} \text{ c.g.s. units} = 1{\cdot}9\times 10^{-2} \text{ c.g.s. units} = 5{\cdot}7 \text{ eV};$$

the electrostatic energy of the ions in the molecule is $E = -\frac{e^2}{r} = 5{\cdot}7$ eV. Consequently the energy of formation of the ionic molecule Na^+Cl^- from neutral atoms is $U_{NaCl} = 4{\cdot}3$ eV $= 101$ kcal./mole. This value agrees well with the dissociation energy of the molecule, $D = 98$ kcal./mole, determined spectroscopically. If we were to take the bond to be covalent, then calculation of the energy of the NaCl molecule as the average of the binding energies in the sodium crystal and in molecular gaseous chlorine,

$$E_{NaCl} = \tfrac{1}{2}\left[U_{Na\ cryst} + \tfrac{1}{2}D_{Cl_2}\right] \approx 37 \text{ kcal./mole}$$

gives a much smaller value.

The ionic character of the bond in the gaseous molecules of alkali-halide compounds is confirmed by the observation that they have constant electric dipole moments. The latter are somewhat nearer to the calculated moments than are the moments obtained for hydrogen halide molecules, in which the bond approximates more closely to the covalent. The dissociation of salts in aqueous solutions with the formation of electrolytes may also serve as an indirect confirmation of the existence of ions. The ionic character of the bond in the NaCl crystal is indicated by data on X-ray scattering intensity.

Stoichiometry, Nonsaturation and Coordination Relationships of the Ionic Bond

Ions of alkali metals and halogens have the same electron configurations as atoms of the adjacent noble gases; that is, they possess close-packed spherically symmetrical electron shells. If the ions were not charged, there would be no chemical interaction between them, just as in the case of atoms of the chemically inert noble gases. The electric charges on the ions bring about their attraction or repulsion and in general determine the stoichiometric composition of the ionic compound, since the sums of the charges on anions and cations must be equal. Among binary ionic compounds the most frequently encountered are compounds of simple stoichiometric compositions AB, AB_2, AB_3, A_2B_3, A_3B_4, examples of which are NaCl, CaF_2, BiF_3, Fe_2O_3, Fe_3O_4 etc.

The interaction of spherically symmetrical ions is undirected and unsaturated. The nonsaturation of the ionic bond is shown by the fact that each ion tends to draw the oppositely charged ions as closely as possible to itself, that is, to form a structure with coordination numbers as large as possible. The possible values of the coordination numbers are determined by the ratios of the ionic radii. Fig. 63 shows the simplest configurations formed by anions of radius r_A around cations of radius r_C. An increase in coordination number is accompanied by an increase in the ratio r_A/r_C. These configurations lose their stability if the cations are so small that they are freely distributed between the anions and can 'wander around' in their respective voids. It can easily be shown that in this case the electrostatic energy of the system of ions will decrease, that is, the stability of the system will grow, if a cation approaches some of the anions surrounding it, or if, in other words, it lowers its coordination number.

If the cation is sufficiently large and separates the anions, then the corresponding configuration will be stable. However, if the ratio of the radii r_C/r_A is such that close contacts are possible between the cations and a large number of anions, then there will be a reconstruction of the configuration in such a way as to increase its coordination number. Table 20 shows the critical ratios of ionic radii for some of the simplest and most symmetrical configurations.

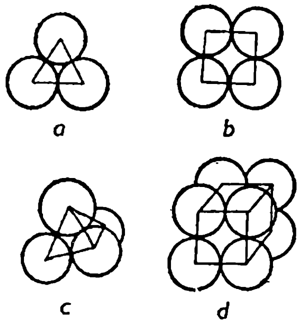


FIG. 63

The simplest configurations in binary molecules and crystals: (*a*) triangle, (*b*) square and octahedron, (*c*) tetrahedron, (*d*) cube.

Table 20

Critical Ratios of the Ionic Radii for Configurations with Different Coordination Numbers *k*

k	Configuration	r_C/r_A
2	dumb-bell	0
3	triangle	$0{\cdot}15 = \sqrt{\frac{4}{3}}-1$
4	tetrahedron	$0{\cdot}22 = \sqrt{\frac{2}{3}}-1$
4	square	$0{\cdot}41 = \sqrt{2}-1$
6	octahedron	$0{\cdot}41 = \sqrt{2}-1$
8	cube	$0{\cdot}73 = \sqrt{3}-1$
12		1·00

The nonsaturation and nondirectivity of the ionic bond lead to the result that in the condensed state, ionic compounds like the majority of substances with metallic bonds, form crystals of the coordination type. Because of the high energy of the ionic bond these crystals have considerable rigidity and high melting points. Many ionic crystals are transparent and diamagnetic, and are insulators. At high temperatures, with the increased mobility of the atoms, there is electrical conductivity of the ionic type, accompanied by a transfer of ions (solid electrolytes).

Experimental Determination of the Electron Density in Heteropolar Compounds

Depending on the differences between the electronegativities of atoms, it is possible to form compounds in which the bonds have greater or less ionic character. In the study of these intermediate types of bond one has to determine the effective charges on the ions arising from the partial redistribution of electron density. Several methods are known for investigating this experimentally (paramagnetic resonance, X-ray spectra, X-ray structural analysis). From the curves of the atomic factors (by extrapolation to the case $\frac{\sin\theta}{\lambda} \to 0$) we can determine the effective electron charges of atoms in crystals. X-ray structural data for the atoms of nickel and nitrogen in nickel phthalocyanine show that $Z^*_{\text{Ni}} = 26$ and $Z^*_{\text{N}} = 7{\cdot}5$, instead of $Z_{\text{Ni}} = 28$ and $Z_{\text{N}} = 7$ for the free atoms. In this compound the metallic atom is bound to four nitrogen atoms, which in turn enter into the more complex framework of the molecule. Experimental data show that part of the electron density of the nickel atom ($\Delta Z \approx 2$ electrons) passes over to the neighbouring, more electronegative, nitrogen atoms, and as a result the charge on each of them is raised by $-\frac{1}{2}$ an electron. In the crystal $HNO_3.3H_2O$, on the other hand, there is a deficit of electrons on the nitrogen atom, $Z^*_{\text{N}} = 5{\cdot}4$, and an excess of them on the oxygen atoms, $Z^*_{\text{O}} = 8{\cdot}5$; this is connected with the greater electronegativity of oxygen atoms compared to nitrogen atoms. For the hydrogen atoms in the water molecule (Fig. 64) we find a slight electron density deficit ($\Delta Z^*_{\text{H}} = 0{\cdot}2e$) and there is an increase in the negative charge on the oxygen atom ($\Delta Z^* \approx 0{\cdot}4e$). Hence it is possible to calculate the dipole moment of the water molecule, $p = 2p_1 \cos \phi/2 = 1{\cdot}2 \times 10^{-18}$ e.s.u. Within the limits of experimental accuracy (the error in the determination of the effective charge on the hydrogen atom is about $0{\cdot}1e$) this value agrees with the directly measured dipole moment of the water molecule, $p = 1{\cdot}84 \times 10^{-18}$ e.s.u.

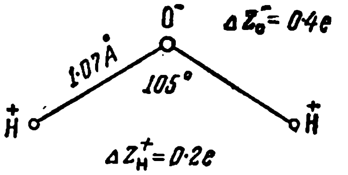


FIG. 64. Electron density distribution in the water molecule.

The Structure of the Molecules As_4S_4 and N_4S_4

The important influence exercised by the difference between the electronegativities of atoms on their bond can be seen by comparing the structures of the molecules of realgar As_4S_4 (Fig. 65) and sulphur nitride. In the realgar molecule the sulphur atoms are arranged at the vertices of a square and form two bonds each with the arsenic atoms, which in their turn form three bonds each (two with S atoms and one with As). The structure of this molecule agrees with the normal valencies of As and S atoms and with the rules for directed valencies, since the bonds are formed by *p*-electrons. The sulphur atoms are more electronegative than the arsenic atoms, so the bonds have a partial ionic character, as is shown in

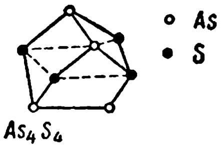


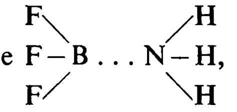
FIG. 65. Structure of realgar and sulphur nitride molecules.

particular by the slight increase in the distance $As^{+\delta}$—$As^{+\delta}$, due to electrostatic repulsion as compared with the distance As—As in the arsenic crystal.

The sulphur nitride molecule has a structure analogous to the realgar molecule, with the sole difference, unexpected at first sight, that the nitrogen atoms have taken the places of the sulphur atoms and the sulphur atoms have replaced the arsenic atoms. In the sulphur nitride molecule there is a special case of isomorphism in which an element of the VI group in the Periodic Table replaces an element of the V group (S⟶As) and, conversely, an element of the V group replaces an element of the VI group (N⟶S). The explanation is that the nitrogen atom is more electronegative than the sulphur atom. The difference between the electronegativities of the atoms leads to a redistribution of the electron density $S^{+\delta}$ and $N^{-\delta}$; this raises the valency of the sulphur atom and lowers that of the nitrogen atom.

The Donor-acceptor Bond

Molecules formed from atoms whose valence shells are filled to differing degrees can form more or less stable molecular compounds with one another. We shall consider the molecules boron trifluoride BF_3 and ammonia NH_3. The BF_3 molecule has a plane triangular configuration (sp^2-hybridisation). One valence level of the boron atom (the acceptor) remains free.

The molecule NH_3 has a pyramidal configuration. In the valence shell of the nitrogen atom (the donor) there is a free electron pair. The molecules BF_3 and NH_3 form together the molecular compound $F_3B.NH_3$, with the structure  $F_3B\ldots NH_3$, in which the boron and nitrogen atoms are joined by a donor-acceptor bond, which makes use of the electron pair from the nitrogen atom and the vacant level of the boron atom. The formation of a molecular compound is accompanied by a marked change in the valence state of the boron atom, with an approximation to tetrahedral sp^3-hybridisation while the BF_3 group acquires a pyramidal configuration. The length of the B—F bond grows to 1·38 Å, which exceeds the length (1·30 Å) of the B—F bond in the molecule BF_3 and approaches the length of the B—F bond in the tetrahedral ion BF_4^-. The length of the donor-acceptor B—N bond (1·60 Å) is considerably greater than the length of the B—N bond in boron nitride BN (1·45 Å). The structure of boron nitride is based on sp^2-hybridisation. Accordingly the strength of the molecular compound $F_3B.NH_3$ is low and amounts to $E \approx 17$ kcal./mole. The donor-acceptor bond is fundamental to the understanding of catalysis processes in chemistry, the properties of semiconductors etc.

Problems

1. Using the Bohr quantisation condition, calculate the internuclear distance, the radius of the molecular orbital and the energy of the molecular hydrogen ion H_2^+.
2. Do the same for the hydrogen molecule H_2. Determine the binding energy of this molecule.

Chapter 5

THE SYMMETRY OF ATOMS, MOLECULES AND CRYSTALS

1. The Symmetry of Finite and Infinite Figures

Symmetry and Symmetry Operations

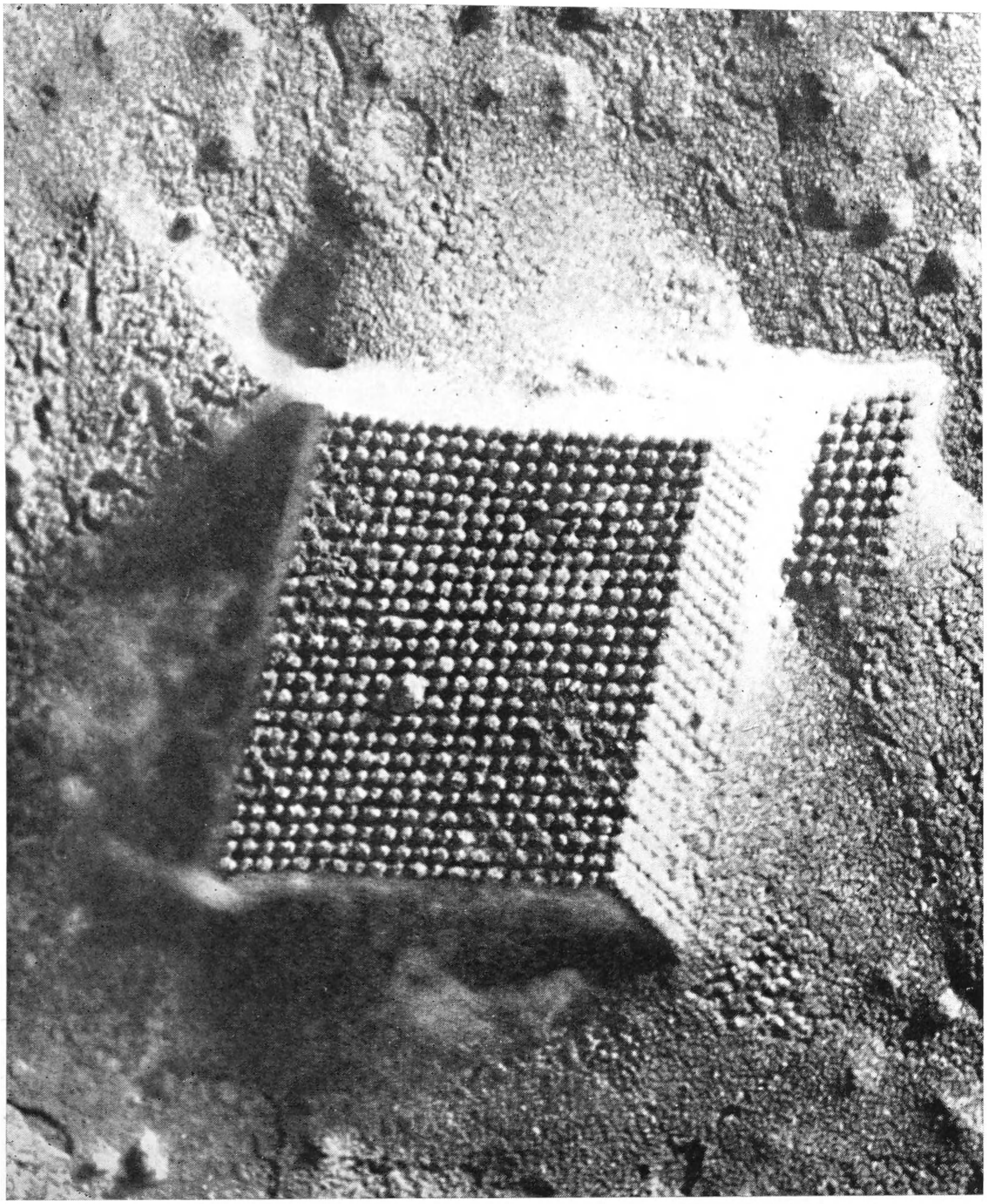
The wave functions of the hydrogen atom, the hybrid wave functions formed from them and the electron density distributions in the atoms all possess symmetry. The symmetry of atoms and the bonds formed by them leads to a definite symmetrical structure in molecules and crystals. A knowledge of the fundamentals of symmetry theory and the methods of representing symmetry is essential for the understanding and description of the structure of molecules and crystals and the analysis of their properties.

The symmetry of any body consists in its property of self-coincidence in different positions in which it may be placed. This property is shown in the spatial repetition of the parts of the body.

Symmetry extends the idea of congruence of figures. Figures may be called congruent if the distances between corresponding points are the same. This condition is necessary, but not sufficient for ordinary congruence, in which the figures coincide on superposition. The condition is observed alike for right-handed and left-handed figures, but these do not coincide with one another when superimposed (for example, right-handed and left-handed gloves, right-handed and left-handed molecules of tartaric acid, right-handed and left-handed quartz). Such figures coincide only after mirror reflection, which transforms a right-handed figure into a left-handed and vice versa. The difference between the right-handed and left-handed forms of molecules and crystals appears not only in their structures but also in certain physical properties, and may be revealed, for example, by a rotation of the plane of polarisation.

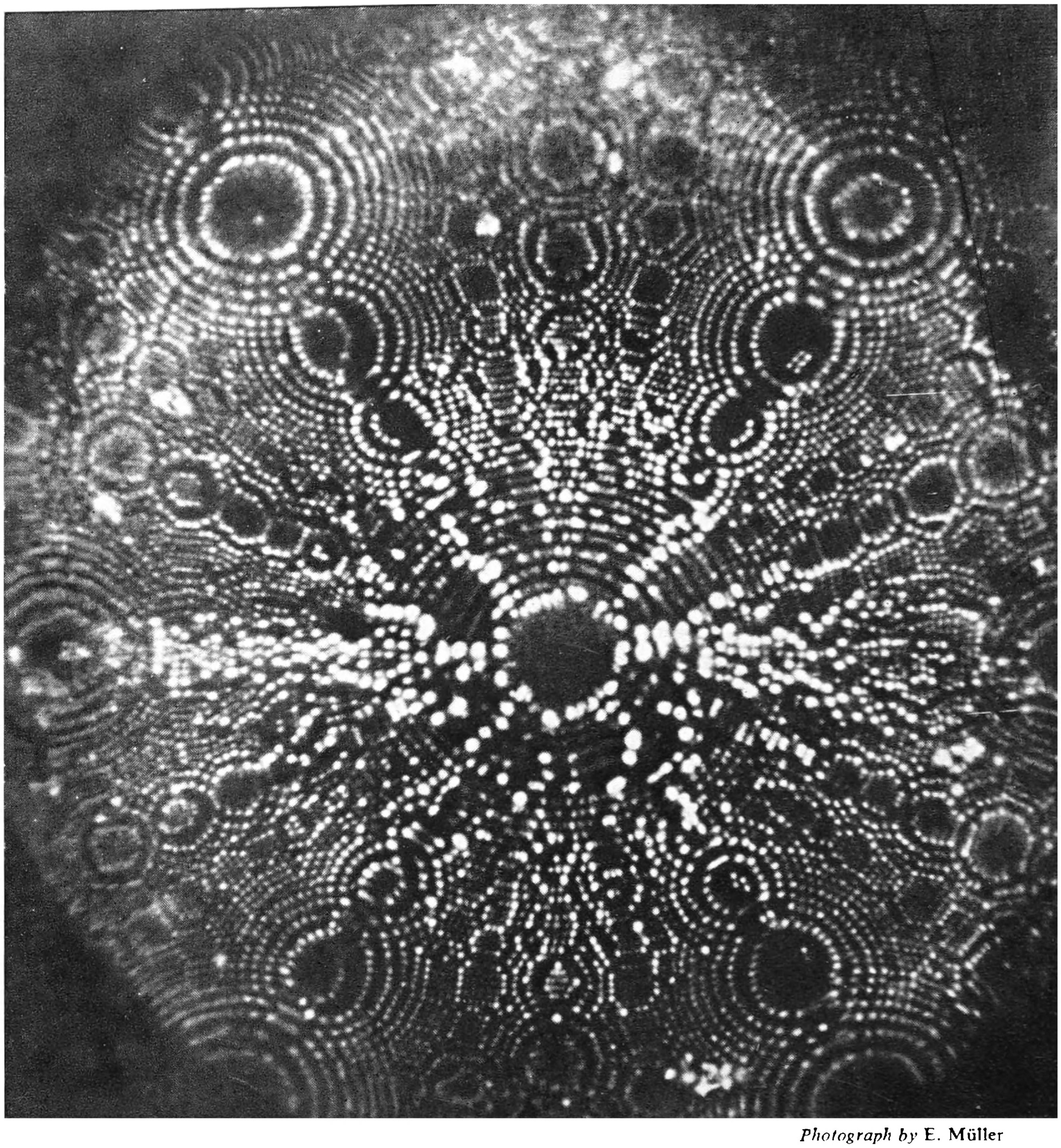
Coincidence of different bodies or self-coincidence of a single body is accomplished by means of *reflections, rotations, inversions and parallel translations*. These are called symmetry operations or symmetry transformations. Symmetry transformations are transformations of coordinates $(x, y, z) \rightarrow (x', y', z')$ without extensions, contractions or shears, in other words such that the distances between points are not altered during the transformation.

The symmetry of crystals can be described in two ways. Firstly, crystals are finite bodies with properties which exhibit *anisotropy*, which is often apparent from their regular form. In this case the crystals may be replaced by polyhedra having the same symmetry as they have, and this symmetry may be described (subject to certain



Photograph by R. W. G. Wyckoff

FIG. 66. Electronmicrograph of a crystal of tobacco mosaic virus.



Photograph by E. Müller

Fig. 67. Field-ion micrograph showing the atomic structure of a spherical crystal of tungsten. The crystal diameter was about 100A.

limitations, to be considered later) in an analogous manner to the symmetry of atoms and molecules and of finite figures in general. Secondly, crystals possess regular periodic structures and they may be looked upon as infinite bodies which are self-coincident under parallel translations which are multiples of the periods of the structure. In this case the symmetry operation is a parallel translation, which does not occur for finite figures.

> The idea that crystals had a regular periodic structure was reached a long time ago, being then based on their macroscopic properties. The study of symmetry provided a sufficiently rigorous basis for the theory of lattice structure and the symmetry law for crystals was established in the first instance by observation. Long before the arrival of accurate methods of investigating and observing the atomic and molecular structure of matter, Federov (1890) developed the theory of space groups, which describes the symmetry of the atomic structures of crystals. A weighty, though indirect, argument for the lattice structure was Laue's discovery (1912) of X-ray diffraction in crystals. The subsequent development of new methods of microscopy has enabled us to observe the molecular and atomic structure and actually see the space lattice. Fig. 66 shows an electron micrograph of a crystal of the tobacco mosaic virus (obtained by Wyckoff); its molecules have a diameter of about 250 Å. The great depth of focus enables us to see the structure in perspective and the photograph graphically reproduces the model of the space lattice. Fig. 67 is a photograph of the surface of a spherical crystal of tungsten with diameter about 100Å, obtained by Müller using his field ion microscope. The high resolving power of this type of projector (2-3 Å) enables us to see the atomic structure of the metal. The spherical crystalline form of the crystal gives a picture resembling the spherical projection of a crystal. The circular and oval contours mark the points of emergence of the normals to the individual planes of the crystal. The central region corresponds to the point of emergence of the normal [011]. The four larger circular regions on the periphery of the photograph represent the points of emergence of the normals [112], $[\bar{1}\bar{1}2]$, $[\bar{1}21]$ and $[1\bar{2}1]$.

The operations of rotation and parallel translation (the latter can be looked on as a rotation through an infinitesimal angle about an infinitely distant axis) do not alter the type of figure. Thus a right-handed figure remains a right-handed figure under these operations. Reflection and inversion alter the type of figure, transforming a right-handed figure into a left-handed. Accordingly symmetry operations may be divided into two groups, those of the 'first' and 'second' kind, depending on whether they do, or do not, alter the type of figure.

The simplest symmetry operation is a reflection. All the other symmetry operations can be reduced to reflections in a series of mirrors. A rotation through an angle α is equivalent to reflections in two planes intersecting along the axis of the rotation, and forming between them an angle $\frac{1}{2}\alpha$. Reflections in a kaleidoscope consisting of intersecting mirrors m_1 and m_2 (Fig. 68*a*) bring about the formation of a series of points $B_1, B_2, B_3, \ldots$, lying on a circle and inclined at angles α to one another.

The parallel translation determined by the vector $\boldsymbol{a}$ is equivalent to reflections in two planes perpendicular to the vector of the translation and a distance $\frac{1}{2}a$ apart.

On successive reflections in the kaleidoscope consisting of the parallel mirrors a rectilinear periodic series of points A_1, A_2, A_3, . . . is formed (Fig. 68*b*).

Inversion in the centre of symmetry is equivalent to reflections in three mutually perpendicular mirrors passing through the centre of symmetry. Compound symmetry operations are formed by a combination of reflections, rotations and inversions, so any symmetry operation, whether of the first or second kind, may be replaced by a number (even or odd respectively) or reflections.

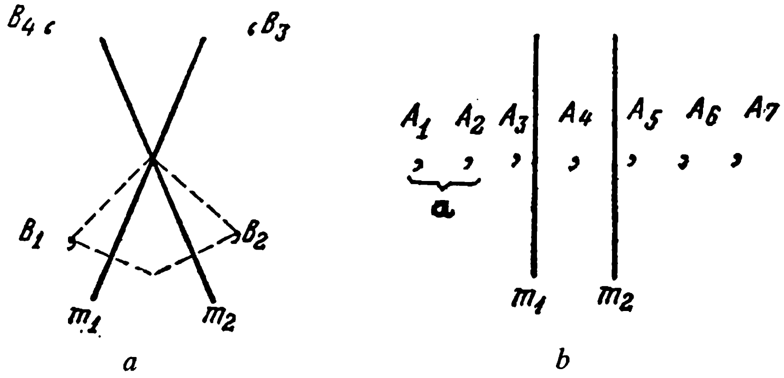


FIG. 68. Translation and rotation as two successive reflections.

The symmetry of any body may be described in terms of reflection operations alone. This was carried out in its application to the symmetry of crystalline forms by the eminent Russian crystallographer Yu. V. Wulf. However for greater clarity and convenience it is better to describe symmetry by use of compound symmetry operations.

Symmetry Elements

Symmetry operations are carried out by means of geometrical elements called symmetry elements. In three-dimensional space these elements are the plane, straight line, point, translation vector and also the more complex symmetry elements arising from combinations of them. The concept of 'symmetry element', like that of 'elementary particle', is relative. All symmetry 'elements', except the mirror plane, are compound. The system of units (elements) in symmetry theory is based on the same principle of uniting units of lowest order to form units of higher order, as the decimal and similar systems of counting. Table 21 gives the most generally used symmetry elements, their symbols, the coordinates of the groups of equivalent points which arise when the symmetry operations prescribed by the symmetry element are performed, and their multiplicities. The multiplicity is the number of equivalent points formed by the symmetry element.

The elementary rotation angle α_n for an n-fold axis of symmetry C is determined by the cyclic condition

$$\alpha_n = \frac{2\pi}{n}, \quad n = 1, 2, 3 \ldots \qquad \ldots [5.1]$$

Any straight line is a one-fold axis of symmetry, C_1. If these are the only one-fold symmetry elements present in a body, it indicates a lack of symmetry. Rotation-inversion axes correspond to a rotation about an axis and reflection in a perpendicular plane. In axes of even order S_{2n} there is neither a separate axis of rotation C_4 nor a symmetry plane perpendicular to it. This can be seen in the example of the crystal pentaerythritol (Fig. 69), which has an axis S_4. The rotation axis of symmetry degenerates into an axis C_n with multiplicity halved. The odd rotation-inversion axes disintegrate into a rotation axis and a plane $S_{2n+1} = C_{2n+1}+m$ perpendicular to it. So only the even rotation-inversion axes S_{2n} are regarded as independent symmetry elements. The group of inversion axes is equivalent to the group of rotation-inversion axes: $S_2 = \bar{1} = i$; $S_4 = \bar{4}$; $S_6 = \bar{3}$.

Table 21

Symmetry Operations, Symmetry Elements and their Symbols

Symmetry operation	Symmetry element	Symbol		Coordinates of equivalent points	Multiplicity (number of equivalent points)
		Schoenflies system	International (Hermann-Maugin) system		
Reflection in a plane	plane of symmetry	C_s	m	xyz; $x\bar{y}z$ (plane m perpendicular to Y-axis)	2
Rotation through angle α_n about an axis	axis of symmetry	C_n	n	xyz; $\bar{x}\bar{y}z$ (axis C_2 coincides with Z-axis)	n
Inversion in a point	centre of symmetry	C_i	$i = \bar{1}$	xyz; $\bar{x}\bar{y}\bar{z}$ (centre of symmetry at origin of coordinates)	2
Rotation about axis and reflection in perpendicular plane	rotation-inversion	S_{2n}			
Rotation about axis and inversion in centre of symmetry, lying on axis	inversion axis		$\bar{n}$		n
Parallel translation	translation vector **a, b, c**			$x = ma$; $y = na$; $z = pa$; m, n, p integers	∞
Reflection in plane and translation parallel to plane	glide plane		a, b, c n, d		∞
Rotation about axis and translation parallel to axis	screw axis		n_m $m = 1, 2, \ldots, n$		∞

Reflection and rotation leave in the same position points lying on a particular plane or straight line respectively. Inversion and operations connected with it (inversion and rotation-inversion axes) leave in the same place a single point of a body. These operations and symmetry elements, all of which leave at least one point of a body unmoved, are possible both in finite and in infinite bodies. The multiplicities of these elements are finite.

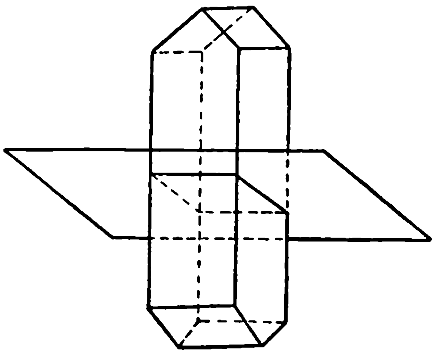


FIG. 69. Rotation-inversion axis S_4 in a crystal of pentaerythritol.

The multiplicities of symmetry operations connected with the parallel translation are infinite. These symmetry operations do not leave a single point of a body unmoved and are possible only in infinite bodies with a *periodic* structure. Successive application of the parallel translation a to the initial point, which we shall call a node, leads to the formation of an infinite, rectilinear and periodic series of nodes, the position of each being determined by the vector of the linear grid, $\boldsymbol{p} = m\boldsymbol{a}$, where $\boldsymbol{a}$ is the length of the axial scale, and m is an arbitrary integer, the coordinate of the nodes in the series. The combination of two translation vectors leads to the formation of a two-dimensional periodic network. The positions of the nodes in a plane network are determined by the binary linear form $\boldsymbol{p} = m\boldsymbol{a}+n\boldsymbol{b}$. The combination of three noncoplanar translation vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ leads to the formation of a three-dimensional infinite and periodic network, called a *space lattice*. The coordinates of the nodes in a space lattice are determined by the trinomial linear form

$$\boldsymbol{p} = m\boldsymbol{a}+n\boldsymbol{b}+p\boldsymbol{c}, \qquad \ldots [5.2]$$

where $[m, n, p]$ are the integral coordinates of the nodes.

The combination of a plane of symmetry with a translation a_n perpendicular to it brings about a translational multiplication of planes of symmetry and the formation of a kaleidoscope of parallel mirrors in which the mirror-symmetrical node is multiplied into a periodic series. When reflection is combined with a parallel translation a new symmetry element appears—the *glide plane*. In Fig. 70 *ab* marks the position of the glide plane; the direction of the displacement is parallel to the plane of the diagram. The glide plane causes the formation of a periodic series A_1, A_2, A_3, $A_4 \ldots$ in the pattern of a 'herring-bone' with period $2a_t$ along the glide plane. The combination of a rotation with a displacement parallel to the axis of the rotation gives a screw

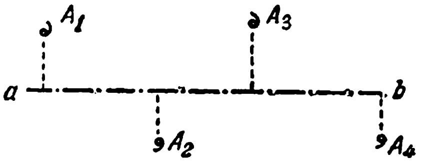


FIG. 70. Transformation in glide plane.

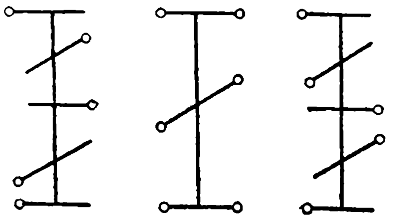


FIG. 71. Group of fourfold screw axes.

motion which corresponds to a new symmetry element—the *screw axis*. The displacements along a screw axis of the nth order are $a_t = (m/n)a$, where $m = 1, 2, \ldots, n$. The total number of screw axes of the nth order is n, including the rotational axis for which $m = 0$ ($m = n$). Fig. 71 shows the group of screw axes of symmetry of 4th order. The axes 4_1 and 4_3 are mirror-symmetrical right-handed and left-handed screw axes. The axis 4_2 contains the rotational axis 2.

2. Symmetry Groups and the Laws of their Formation

The Symmetry Group as a Closed Set

Combinations of symmetry elements generate new symmetry elements, in as much as the corresponding symmetrical transformations may be applied both to the points of a body and also to symmetry elements. The combination of two symmetry elements $A+B$ generates at least one other symmetry element C, which belongs to the group $A+B+C$. If the multiplicities of the elements A and B are greater than two, then, not one, but several new symmetry elements C_1, $C_2 \ldots$ arise. The performance of the symmetry transformations prescribed by these elements will lead to the multiplication of the initial elements, $A_1, A_2, \ldots$ and $B_1, B_2 \ldots$

If no limitations are imposed on the initial set of symmetry elements $A+B$, then the continuation of the symmetry transformations with the newly arising elements brings about a chain process in which the number of symmetry elements grows with a high multiplication coefficient. As a result the symmetry elements in a body may be distributed as densely as we please. This is possible in only one case—that of an *isotropic* body, which possesses the highest form of symmetry—spherical symmetry. In order to obtain other forms of symmetry it is essential that the symmetry elements be combined, not arbitrarily, but in accordance with strict rules. The basic condition for such a combination is that it must give rise to a group, that is, to a set of symmetry elements closed for all symmetrical transformations prescribed by the elements belonging to the group.

The Combination of Rotation Axes and Planes of Symmetry

To any rotation axis of symmetry C_n, with direction which we shall take as vertical, we may annex a *horizontal* plane of symmetry h perpendicular to it. Reflection in the plane of symmetry will convert the polar axis of symmetry C_n into a bilateral nonpolar axis denoted by the symbol C_{nh}. Polar axes of symmetry are characteristic of unclosed crystalline forms in the shape of pyramids (three-faced, square etc.). Nonpolar axes of the type C_{nh} are characteristic of closed forms in the shape of bipyramids.

Across a rotation axis of symmetry C_n we may draw a *vertical* plane of symmetry v. Rotations about the axis form a group consisting of n equivalent planes of symmetry intersecting on the axis of symmetry. A second family of planes of symmetry, also parallel to the axis of symmetry, passes through the bisectors of the angles between members of the first family of planes. This combination of an axis and planes of symmetry is denoted by the symbol C_{nv}. In the particular case C_{3n} these two families of planes of symmetry coincide with one another.

The combination of axes of symmetry with *oblique* planes of symmetry will produce equivalent intersecting axes of symmetry. This suggests that we should consider possible combinations of axes of symmetry.

Combinations of 2-Fold Axes of Symmetry. Axes of Symmetry of Regular Polygons

Consider two intersecting axes of symmetry C_2 forming an angle β. Draw a plane P through the axis of symmetry. Transforming around the twofold axes of symmetry produces new axes lying in the plane P and forming angles between them which are multiples of β. From the cyclic condition it follows that $n\beta = 2\pi$, i.e. the twofold axes form a group of axes of symmetry of a regular n-gon. To the group there also belongs the perpendicular to the twofold axes of symmetry C_n. This combination of axes of symmetry is denoted by the symbol D_n.

Combinations of Higher Axes of Symmetry. Axes of Symmetry of Regular Polygons

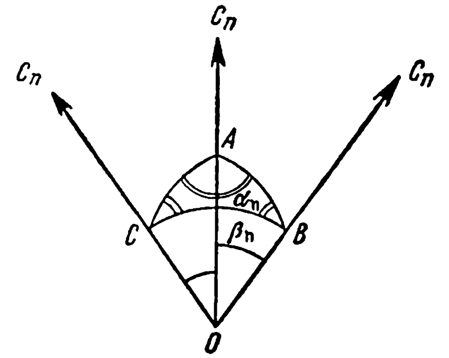


FIG. 72. The group of equivalent axes of symmetry C_n.

Let OA and OB (Fig. 72) be two axes of symmetry C_n, forming the least possible angle β_n. A rotation through an angle α_n about the axis OA transforms the axis OB into the axis OC. A single rotation about the axis OC must transform the axis OA into OB. Consequently all the angles are α_n and the sides β_n in the spherical triangle ABC. From the formula for a spherical triangle we obtain

$$\begin{aligned}\cos\beta_n &= \cot\alpha_n \cot\tfrac{1}{2}\alpha_n = 1/(2\sin^2\tfrac{1}{2}\alpha_n)-1\\ &= 1/(2\sin^2\pi/n)-1. \qquad \ldots [5.3]\end{aligned}$$

Formula [3] indicates the possible cases in which higher axes of symmetry may be combined. As the multiplicity n increases, the denominator in formula [3] becomes

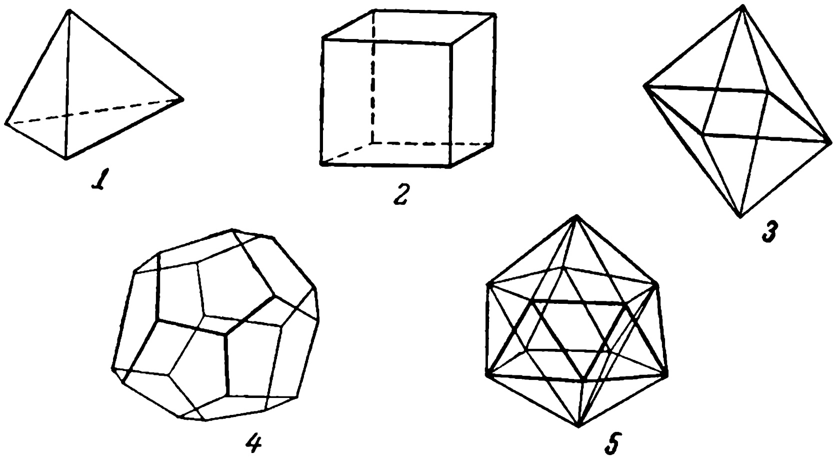


FIG. 73. The regular polyhedra: 1, tetrahedron; 2, cube; 3, octahedron; 4, pentagonal dodecahedron; 5, icosahedron.

smaller and the value of $\cos\beta_n$ greater. The limiting value $\cos\beta_n = 1$, $\beta_n = 0$, corresponding to coincidence of the combining axes of symmetry, is attained for $n = 6$. This trivial case is possible for any axes of symmetry. Nontrivial combinations of higher axes of symmetry ($\beta_n \neq 0$) are possible for three types of axes, shown in Table 22.

These combinations correspond to the arrangements of the axes of symmetry of the regular polyhedra. The faces of regular polyhedra are regular and congruent polygons; the plane angles at their vertices are equal. There are in all 5 regular polyhedra (Fig. 73), whose description is given in Table 23.

The number of faces F, of edges E and vertices V satisfy Euler's formula

$$F-E+V = 2. \qquad \ldots [5.4]$$

Table 22

Combination of Higher Axes of Symmetry

n	$\cos\beta_n$	β_n	Characteristic arrangement of axes
3	$-\frac{1}{3}$	109° 28′	group of 3-fold axes of symmetry of a tetrahedron
4	0	90°	group of 4-fold axes of symmetry of a cube
5	0·263	74° 50′	group of 5-fold axes of symmetry of an icosahedron

Table 23

Description of Regular Polyhedra

Name	Symmetry	Form of faces	F	E	V
Tetrahedron	T	triangle	4	6	4
Cube	O	square	6	12	8
Octahedron	O	triangle	8	12	6
Pentagonal dodecahedron	J	pentagon	12	30	20
Icosahedron	J	triangle	20	30	12

The cube and the octahedron possess the same point symmetry, denoted by the symbol O. The complete group of axes of symmetry comprises the 4-fold, 3-fold and 2-fold $3C_4\ 4C_3\ 6C_2$. The symmetry group of the tetrahedron $T\ 3C_2\ 4C_3$ is a subgroup of the group O. The pentagonal dodecahedron and the icosahedron also possess the same symmetry, denoted by the symbol J. In the group J there are mutual combinations of fivefold, threefold and twofold axes of symmetry $6C_5\ 10C_3\ 15C_2$. The groups T, O and J exhaust the three possible cases of mutual combination of higher axes of symmetry.

Combinations of Rotation Axes with Perpendicular Translations.
Fundamental Symmetry Law of Periodic Structures

The combination of a rotation with a translation parallel to the axis of the rotation, which we considered earlier, does not impose any limitations on the multi-

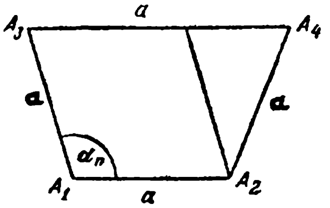


FIG. 74. Combination of a rotation axis with a perpendicular translation.

plicity of the axis of rotation. On the other hand, the combination of a rotation with perpendicular translations does impose important limitations.

Fig. 74 shows the start of the process of multiplication of a rotation axis C_n, perpendicular to the plane of the diagram and initially in position A_1. The translation $\boldsymbol{a}$ gives the axis A_2. Rotations about the axes A_1 and A_2 generate new axes A_3 and A_4, lying on a straight line parallel to the nodal series A_1 and A_2. The distance between the points A_3 and A_4, $a(1+2\cos\alpha_n)$, must be a multiple of the length of the translation vector. Hence we obtain the equation

$$2\cos\alpha_n = p, \qquad \ldots [5.5]$$

where p is an integer. Equation [5] has 5 solutions ($\alpha = 0, 180, 120, 90, 60°$). If we discount the trivial axis C_1, there are only 4 possible types of rotation axes in periodic structures—C_2, C_3, C_4 and C_6. An illustration of this theorem can be found in parquet patterns. Continuous covering of a plane with identical parallel polygons is possible only with parallelograms, rectangles, squares, regular rhombuses and hexagons.

The space lattice represents a set of equal, parallel and periodically arranged two-dimensional lattices. So in crystals only the 4 previously indicated types of axes of symmetry are possible. This conclusion, which contains the gist of the fundamental law of symmetry in crystals, is confirmed by all the experimental crystallographic material which has been accumulated in the course of several centuries.

Point Groups and Space Groups

Symmetry groups containing only the operations of reflection, rotation and inversion, and not containing any translations, are called point groups. These groups leave at least one point of a body unmoved and describe the symmetry of finite figures: atoms, molecules, crystalline polyhedra and other symmetrical bodies. They also describe the symmetry of continuous anisotropic media.

Symmetry groups containing parallel translations describe the symmetry of infinite bodies with a periodic structure. The most important for us are *plane* and *space* groups, which contain two and three independent translations respectively. As a result of the application of the translation, each symmetry element is multiplied into

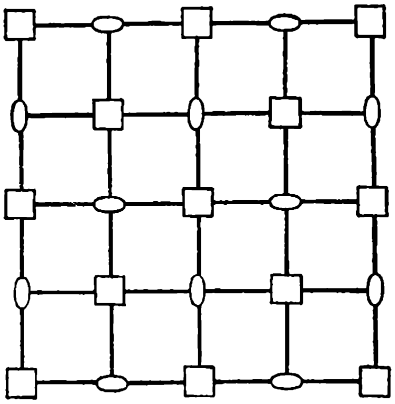


FIG. 75. Combination of fourfold axes of symmetry with perpendicular translations.

a family of identical and parallel elements. Fig. 75 shows a set of fourfold axes with a perpendicular translation; this leads to the formation of two families of parallel fourfold axes and two families of twofold axes parallel to them. The set of these symmetry elements constitutes one of the 17 plane symmetry groups. These groups describe the symmetry of the faces of a crystal, and also the symmetry of the projection of its structure onto one of its faces.

Space groups characterise the symmetry of three-dimensional periodic structures, including the structures of crystals. The structure of a crystal may be obtained if its space group is known and the positions of its particles are given relative to the symmetry elements of the group for one independent region. By means of the symmetry operations determined by the space group the independent region is multiplied infinitely and fills space completely. Within each space group it is possible to distinguish a *subgroup of parallel translations*, which is characterised by a trio of independent translation vectors forming in conjunction with each other a space lattice. In this process the whole crystal is subdivided into a series of identical and parallel parallelepipeds—its unit cells. The elementary parallelepipeds are formed by adding together a certain number of independent regions transformable into one another by means of reflections, rotations and inversions.

Combinations of symmetry elements of space groups are analogous to combinations of symmetry elements of the corresponding point groups. If we suppose that the translation vectors in the space group are null, that is, if we produce its symmetry elements across the origin of coordinates, we obtain one of the point symmetry groups.

3. The Symmetry of Atoms and Molecules

Point Groups

Let us consider point groups, that is, possible combinations of reflections, rotations and inversions.

1. *Groups* C_n. Each group contains only one symmetry element—an n-fold rotation axis. The group C_1 does not have any symmetry elements.

2. *Groups* S_{2n}. Each group has only one rotation—inversion axis of even order. The group $S_2 = C_i = i$ has only a centre of symmetry. Centres of symmetry are also possessed by groups with $2n = 4p+2$, where p is an integer.

3. *Groups* C_{nh}. Each group contains a nonpolar axis C_n and a perpendicular plane of symmetry. The group $C_{1h} = C_s = m$ contains only one plane of symmetry. Groups with even n have a centre of symmetry.

4. *Groups* C_{nv}. Each group contains an n-fold axis and n planes of symmetry passing through it. The group $C_{1v} = C_{1h} = C_s = m$ has only a plane of symmetry.

5. *Groups* D_n. Each group contains an axis of the n-th order and n double axes of symmetry perpendicular to it. The group $D_1 = C_2$. The group D_2 consists of three mutually perpendicular nonequivalent twofold axes.

6. *Groups* D_{nh} are obtained by adding a horizontal plane of symmetry passing through the twofold axes. The combination of a twofold axis with a plane of symmetry gives rise to a perpendicular plane of symmetry. So the group D_{nh}, apart from the horizontal plane, has n vertical planes of symmetry. Even groups with $n = 2p$ possess a centre of symmetry.

7. *Groups* D_{nd} with vertical diagonal planes of symmetry passing between the

twofold axes. For odd n, the axes of symmetry C_2 are perpendicular to the planes of symmetry, so the corresponding groups possess a centre of symmetry.

The following groups are the symmetry groups of the regular polyhedra, and contain only one representative each.

8. *The group T*—the axial group, consisting of the axes of symmetry of the regular tetrahedron: $3C_2\ 4C_3$. The twofold axes are mutually perpendicular; the threefold axes form with each other tetrahedral angles: $\phi = 109° \ 28'$.

9. *The group O*—the axial group consisting of the axes of symmetry of the regular octahedron or cube: $3C_4\ 4C_3\ 6C_2$.

10. *The group T_d*—the complete symmetry group of the tetrahedron; it is obtained by adding to the group T, 6 planes of symmetry passing through an axis C_2 and two axes C_3; axes C_2 are transformed into axes S_4.

11. *The group T_h* is obtained by adding to the group T three mutually perpendicular planes of symmetry passing through two axes C_2; axes C_3 are transformed into axes S_6.

12. *The group O_h* is obtained from the group O by adding a centre of symmetry and represents the complete symmetry group of the octahedron or cube; axes C_3 are transformed into axes S_6.

13. *The group J*—the axial group containing the axes of symmetry of the regular icosahedron.

14. *The group J_h*—the complete symmetry group of the icosahedron, obtained by adding a centre of symmetry to the group J.

Regular Systems of Points (Complexes)

With a given point group we may associate in a rational manner a chosen system of coordinates. We shall take an arbitrary point with coordinates (x, y, z) in general position, that is, a point not lying on any of the symmetry elements. The symmetry transformations of the group multiply this point, and we obtain a group of N equivalent points forming a complex or regular system of points. By using the symmetry transformations, it is not difficult to express the coordinates of all these points in terms of the coordinates of the initial point. The number N is called the multiplicity of the general position.

Multiplicities of General Positions of Point Groups

Group	C_n, S_n	C_{nh}, C_{nv}, D_n	D_{nh}, D_{nd}	T	O, T_d, T_h	O_h	J	J_h
Multiplicity N	n	$2n$	$4n$	12	24	48	60	120

If the initial point is taken in a particular position, that is, on a symmetry element with multiplicity n_i, then its own multiplicity is diminished. The multiplicity of a particular position is $n' = N/n_i$.

If instead of a point in general position we take a plane lying obliquely in relation to the symmetry elements, then we obtain a set of equivalent planes forming a figure (polyhedron). The multiplicities of the figures are the same as the multiplicities of the complexes. Fig. 76 shows the general figure of the group O_h, which is a 48-hedron, formed from identical triangular faces. This polyhedron can be obtained from particular figures, for example from a cube if we divide each face into 8 sections, or from an octahedron by dividing each face into 6 sections. Stereographic projection is often used for the representation of point symmetry. Fig. 77 represents the stereographic projection of the point group O_h. The nine planes of symmetry dissect the sphere into 48 identical spherical triangles. The intersections of the planes are the axes of symmetry of the 4th (square), 3rd (triangle) and 2nd (oval) orders. The crosses (points on the upper hemisphere) and circles (points on the lower hemisphere) represent the projections of the 48 equivalent directions. These directions may be regarded as the normals to the planes of the general figure, or as straight lines drawn through the points of the general complex.

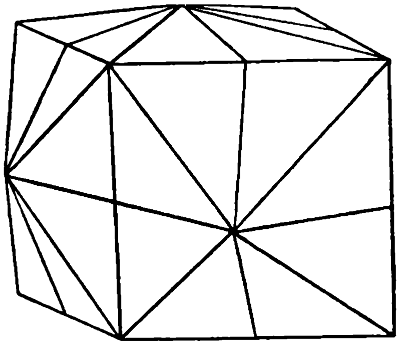


FIG. 76. General figure of the point group O_h.

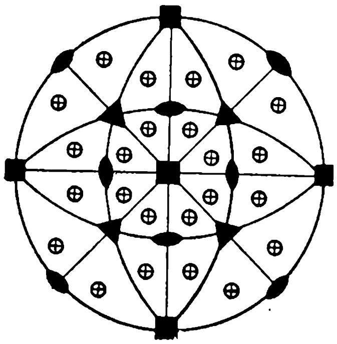


FIG. 77. Stereographic projection of the point group O_h.

4. The Symmetry of Crystals

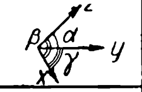
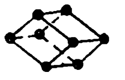
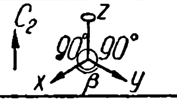
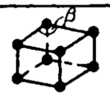
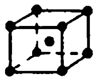
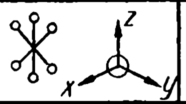
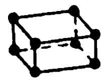
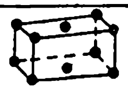
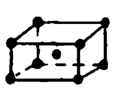
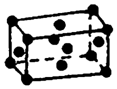
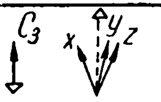
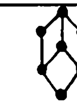
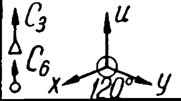
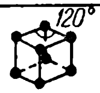
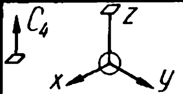
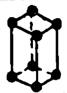
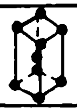
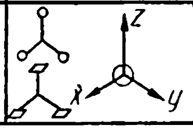
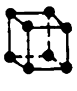
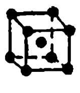
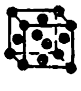
The Crystal as an Anisotropic Continuum and as a Homogeneous Discontinuum

A crystal is a discontinuous medium with a triply periodic distribution of matter (homogeneous discontinuum). All the properties of a crystal are determined by its composition and structure. However in many cases a crystal may be regarded as an unbroken continuous vector medium whose properties depend on direction (anisotropic continuum). The properties of a crystal are determined by the cumulative influence of a sufficiently large number of its constituent particles. For example, the density of a crystal, as determined by weighing, represents the density averaged over the unit cells, no account being taken of the discreteness of the density distribution within the elementary cell, although it is an immediate consequence of the X-ray structural data. The macroscopically measured thermal expansion of a crystal is the averaged resultant of the relative displacements of the equilibrium positions of the atoms caused by the variation in temperature.

A knowledge of the atomic structure is not essential in order to measure the thermal expansion tensor; however, it is absolutely necessary in order to understand the properties of these tensors in individual concrete cases.

Some properties of a crystal, such as for example its density, are scalar. The majority of the other properties, for instance thermal expansion (in non-cubic crystals), velocity of sound, elastic moduli, strength, electrical conductivity etc. are vector properties and are represented by vector models. The length of a radius vector in such a model gives the magnitude of the property when measured in the direction of that radius vector. The symmetry of the vector model is described by one of the point

Table 24
Crystal Systems and Bravais Lattices

N	Crystal system	Principal axis of symmetry	Abbreviation	Arrangement of the axes of symmetry and directions of co-ordinate axes	Unit cell parameters and angles	Bravais lattice			
							complex		
						Simple (primitive) $[[000]]$ $z=1$	base-centred $[[000;\frac{1}{2}\frac{1}{2}0]]$ $z=2$	body-centred $[[000;\frac{1}{2}\frac{1}{2}\frac{1}{2}]]$ $z=2$	face-centred $[[000;\frac{1}{2}\frac{1}{2}0;\frac{1}{2}0\frac{1}{2};0\frac{1}{2}\frac{1}{2}]]$ $z=4$
1	Triclinic	C_1	*tr*		irregular parallelepiped $a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	 Γ_{tr}	—	—	—
2	Monoclinic	C_2	*m*		right prism on a parallelogram as base $a \neq b \neq c$ $\alpha = \gamma = 90° \neq \beta$	 Γ_m	—	 Γ_m'	
3	Rhombic	$3C_2$ *	*o*		right angled parallelepiped $a \neq b \neq c$ $\alpha = \beta = \gamma \neq 90°$	 Γ_o	 Γ_o'	 Γ_o''	 Γ_o'
4	Rhombohedral	C_3	*rh*		rhombohedron $a = b = c$ $\alpha = \beta = \gamma \neq 90°$	 Γ_{rh}	—	—	—
5	Hexagonal	C_3 ** or C_6	*h*		right prism on a rhombus as base $a = b \neq c$; $\alpha = \beta = 90°$ $\gamma = 120°$	 Γ_h	—	—	—
6	Tetragonal	C_4	*t*		rectangular prism $a = b \neq c$; $\alpha = \beta = \gamma = 90°$	 Γ_t	—	 Γ_t'	
7	Cubic	$3C_2$ *** or $3C_4$	*c*		cube $a = b = c$ $\alpha = \beta = \gamma = 90°$	 Γ_c	—	 Γ_c'	 Γ_c''

* The symmetry axis are not equivalent. ** The difference from rhombohedral crystals is that here the basic translations are not perpendicular to the principal axis. *** The symmetry axis are equivalent.

groups and may be greater, but never less, than that of the point symmetry group of the crystal, since it also depends on the character of the property being measured. The symmetry of a crystal as an anisotropic continuum is described by a point group, while its symmetry as a homogeneous discontinuum is described by a space group.

Systems and Crystal Classes

The description of the geometry and properties of crystals is considerably simplified if we take the symmetry of the crystal into consideration when choosing our system of coordinates. This means that the axes of coordinates are chosen parallel or perpendicular to the principal axes and planes of symmetry. From the symmetry law for crystals it follows that there are six types of crystallographic coordinate system. Table 24 gives the descriptions of these systems.

In Table 24 C_n is the symbol for an n-fold axis of symmetry, a, b, c are the lengths of the edges, α, β, γ are the angles of the unit cell opposite to them and Z is the number of particles in unit cell.

> If in choosing the coordinate axes we limit ourselves to considering the properties of the crystal as a continuous medium (anisotropic continuum), then we cannot distinguish rhombohedral crystals from hexagonal and they are described by a single crystal coordinate system. These crystals do however differ in their type of space lattice, and this can be established by studying the crystal as a discontinuous medium (homogeneous discontinuum). The Bravais lattice for rhombohedral crystals is a rhombohedron, and for hexagonal ones a hexagonal prism. These Bravais lattices can be distinguished by the extinction laws established by structural methods. The indices of reflection, hkl, of the rhombohedral crystal satisfy the well-known rhombohedron condition $h-k+l = 3n$, $n = 0, \pm 1, \pm 2, \ldots$

The point groups detailed in §3 describe the symmetries of atoms, molecules, regular polyhedra etc. The study of the symmetry of crystals usually begins with a study of the symmetry of their external shapes, which is described by one of the point groups. The fact that there are so few types of axes of symmetry in crystals limits the number of point groups that are possible. From the first seven types of point groups (§3), which contain an infinite number of members, we choose those groups for which n takes one of the following values: 2, 3, 4 and 6. This gives 27 groups, which are shown in the first six rows of Table 25.

From the remaining seven groups, which are each the sole representative of its type, we choose 5 groups (Nos. 8-12) related to the cubic system. The groups J and J_h, which contain 5-fold axes, are impossible in crystals. Thus the total number of crystalline classes is 32. The derivation of the crystal symmetry classes was given by Hessel (1836), but passed unnoticed in the literature. The complete derivation of crystalline classes was given independently of Hessel by the great Russian engineer and crystallographer A. V. Gadolin (1876).

Unit Cells

The subgroup of parallel translations subdivides the crystal into a set of equal parallelepipedal unit cells the edges of which are the three basic translation vectors.

Table 25
Crystal Classes and Systems

System	Crystal classes†						
Triclinic	$C_1 = 1$	$C_i = \bar{1}$					
Monoclinic	$C_2 = 2$	$C_s = m = \bar{2}$	$C_{2h} = 2/m$				
Rhombic					$C_{2v} = 2mm$	$D_2 = 222$	$D_{2h} = mmm$
Rhombohedral	$C_3 = 3$	$C_{3i} = \bar{3}$			$C_{3v} = 3mm$	$D_3 = 32$	$C_{3d} = \bar{3}m$
Tetragonal	$C_4 = 4$	$S_4 = \bar{4}$	$C_{4h} = 4/m$	$D_{2d} = \bar{4}2m$	$C_{4v} = 4mm$	$D_4 = 42$	$D_{4h} = 4/mmm$
Hexagonal	$C_6 = 6$	$C_{3h} = \bar{6}$	$C_{6h} = 6/m$	$D_{3h} = \bar{6}m$	$C_{6v} = 6mm$	$D_6 = 62$	$D_{6h} = 6/mmm$
Cubic	$T = 23$		$T_h = m3$	$Td = \bar{4}3m$		$0 = 43$	$0_h = m3m$

† Crystal classes are given in Schoenflies and international (Hermann-Mauguin) notations. The latter notation gives the set of symmetry elements which determine the point group. An oblique stroke denotes normality to the corresponding symmetry elements.

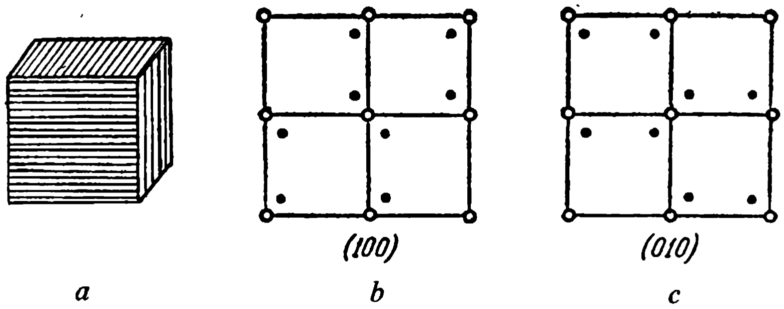


FIG. 78. Symmetry and structure of the pyrites crystal.

Corresponding to the seven systems, there are seven types of simple unit cells, differing in their symmetry. Table 24 gives the shapes of the unit cells and the ratios between the lengths of the translation vectors (lattice periods). Cubic symmetry is found not only in crystals with three fourfold axes, but also in those with three equivalent twofold axes. In the latter case three mutually perpendicular twofold axes are chosen as the axes of coordinates. An example of such a crystal is pyrites FeS_2, which crystallises in the form of cubes (Fig. 78*a*). As the structure of pyrites shows (Fig. 78*b*, *c*), there are twofold (screw) axes passing perpendicularly across the faces of the cube. The lowering of the symmetry in the crystals is shown by the etch figures which on the cube faces of a pyrites crystal are striations parallel to the edges of the cube.

Simple Treatment of Monoclinic and Triclinic Crystals

In crystals of high and medium symmetry the coordinate axes are chosen simply to suit the symmetry of the crystals. In crystals of lower symmetry (monoclinic and triclinic) there is some freedom of choice of axes which is greatest in the triclinic system. In the triclinic lattice the unit parallelepiped can be chosen in an infinite number of ways, with the sole condition $v = v_{\min} = 1$, where v is the volume of the elementary parallelepiped, and $v_{\min} = 1$ is the volume associated with 1 node of the lattice.

B. N. de Launay gave the analytical algorithm for reducing the elementary parallelepipeds to the basic form. By means of successive transformations of the axial vectors of the lattice we can arrive at a unique triad of vectors. This solves the problem of choosing a single system of coordinates to describe a given lattice. This unit cell approximates as closely as possible to a cell in which the diagonals are so far as possible not shorter than the sides and the coordination angles approximate to right angles.

For monoclinic crystals, which in practice are encountered experimentally much

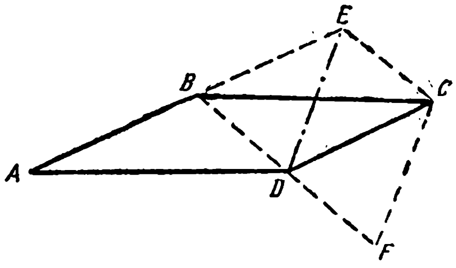


FIG. 79. Transformations of monoclinic unit cell.

more frequently than triclinic, the *b* axis is readily established and the problem reduces to the simple choice of a unit parallelogram in the monoclinic plane. In this case it is possible to use a simple geometrical reduction algorithm which consists in comparing the lengths of the diagonals and the edges of an elementary parallelogram. If we first find the parallelogram *ABCD* (Fig. 79), in which the diagonal *BD* is shorter than the sides, then on carrying out the transformation mentioned we obtain successively the elementary parallelograms $ABCD \rightarrow BDCE \rightarrow CEDF$. The last is the basic parallelogram, since its sides are shorter than its diagonals.

Subgroups of Parallel Translations. Bravais Lattices

In each space group we can distinguish the subgroups of parallel translations. This enables us to consider the set of all possible translational symmetry subgroups independently of the rotational subgroups. This was first done by Bravais in the middle of the last century. Apart from the seven simple lattices the cells of which are shown in Table 24 and contain one node each, it is possible to form complex lattices by either body-centring or face-centring the cells, there being 2 or more nodes per cell. A consideration of the possible cases of cell-centring shows that there are in all 14 different types of translational lattices; they are known as Bravais lattices. These 14 types of lattices, shown in Table 24, represent the set of subgroups of parallel translations which are possible in space groups.

In the cubic system, apart from the simple cubic lattice, body-centred cubic (b.c.c.) and face-centred cubic (f.c.c.) lattices are also possible; they contain 2 and 4 nodes per cell respectively. It is not possible to have a cube with only one pair of faces centred, since this would lead to a lowering of the symmetry and transform the cubic cell into a tetragonal one. In a complex Bravais lattice it is always possible to choose the base vectors in such a way that the new cell contains only one particle. Formally such a cell belongs to a lower symmetry class, which disguises the true symmetry of the crystal.

Fig. 80 shows the choice of a primitive rhombohedron with angle $\alpha_{rh} = 60°$ in an f.c.c. cell. If structural investigations reveal the presence of a rhombohedral cell with angle $\alpha_{rh} = 60°$, then one should make sure whether the crystal is cubic, since then it is convenient to transform the coordinate axes and describe it as f.c.c.

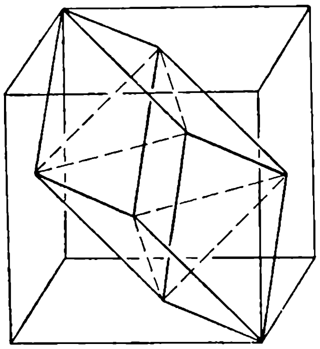


FIG. 80. Primitive rhombohedron in a cubic face-centred cell.

Space (Federov) Groups

The derivation of all possible space groups is considerably more complicated than that of the point groups and symmetry classes in crystals. This derivation was first given by Fedorov (1890). The total number of space groups is 230. The idea behind the derivation is based on the connection between space and point groups. For any crystal class we consider in succession the sets of symmetry elements in that class with translational subgroups which are possible in the system to which the given crystal class is related. As a result we obtain the space groups related to that symmetry class.

For example, the point group C_2-2 of the monoclinic system contains one symmetry element—the axis of symmetry 2, the combination of which with a translation parallel to it leads to the formation of the screw axis of symmetry 2_1. The combination of twofold axes of symmetry with translations perpendicular to them leads to two space groups $P2$ and $P2_1$ with primitive cells containing only rotational or only screw axes. The combination of twofold axes with a translation inclined obliquely to them gives rise to the space group $C2$ with a cell centred in a single pair of faces and containing both types of axes of symmetry, 2 and 2_1. The three space groups $P2$, $P2_1$ and $C2$ exhaust the set of space groups of class 2.

Table 26 shows the distribution of space groups by crystal classes.

Table 26

Distribution of Space Groups over Crystal Classes

Class	Number of groups	Class	Number of groups	Class	Number of groups	Class	Number of groups
1	1	4	6	$\bar{3}$	2	$6mm$	4
$\bar{1}$	1	$\bar{4}$	2	$3mm$	6	62	6
m	4	$4/m$	6	32	7	$6/mmm$	4
2	3	$\bar{4}2m$	12	$\bar{3}m$	6	23	5
$2/m$	6	$4mm$	12	6	6	$m3$	7
$2mm$	22	42	10	$\bar{6}$	1	$\bar{4}3m$	6
222	9	$4/mmm$	20	$6/m$	2	43	8
mmm	28	3	4	$\bar{6}m$	4	$m3m$	10

11 pairs of space groups differ only in the directions of rotation of the screw axes. Dextro- and laevo-rotatory crystals can only belong to these groups. The X-ray extinctions enable us to distinguish 120 space groups. If we take into account the ratios of the period lengths, the number of space groups that are distinguishable by X-ray spectroscopy goes up to 122.

Distribution of Crystals by Space Groups

A statistical study of the results of X-ray spectrographic determination of the space groups for ~4000 crystals where the structures have been sufficiently well investigated reveals a number of interesting laws. We may note especially the extraordinary inequality of distribution of the crystals among the 219 (not counting the 11 enantiomorphic) space groups. For 53 groups not one crystal has yet been found, 32 groups have a single representative each, and 20 two. Thus 105 Fedorov groups (that is, almost half) account for only about 0·2% of the crystals investigated, or 2½ times less than the share of a single Fedorov group if the distribution were uniform.

An analysis of the distribution of crystals according to their numbers within each space group shows that the selective choice of this or that space group is governed by general principles of crystal chemistry. Many inorganic coordination crystals with the ionic or metallic type of bond crystallise in the more symmetrical space groups related to the cubic and hexagonal systems, whereas organic molecular crystals in the majority of cases crystallise in the less symmetrical space groups related to the monoclinic and rhombic systems.

The crystal structure of any substance is determined in the last analysis by thermodynamic factors. Out of the innumerable varieties of conceivable structural types a given substance in equilibrium conditions 'chooses' that type which corresponds to its minimum of free energy. For crystals with nondirected bonds the thermodynamic minimum free energy condition can in many cases be replaced with sufficient accuracy by the geometric condition of minimum volume, that is, by using the principle of closest packing. These cases provide the simplest way of establishing the connection between Fedorov's geometrical theory of crystal structure and Gibbs' thermodynamic theory.

The structures of inorganic crystals with ionic and metallic bonds can be represented very conveniently by close packing of spheres. The number of close packed systems is infinitely large but Belov showed that they are all related to only eight space groups, belonging to the cubic, hexagonal and rhombohedral systems: $C3m$, $R3m$, $C6m2$, $C6mc$, $C\bar{3}m$, $R\bar{3}m$, $C\frac{6}{m}mc$, $Fm3m$. These space groups contain the subgroup $C3m$, which is the group of minimum symmetry for any close packed system. The cubic group $Fm3m$ accounts for about 10% of all crystals whose structures have been determined, and the hexagonal group $C\frac{6}{m}mc$ for 4·3%. Inorganic crystals, including minerals, often crystallise also in the rhombic group *Pnma*. The explanation is that in this group it is possible to have close packing with a slight distortion which reduces its symmetry to rhombic.

In molecular crystals we have to consider packings of molecules with shapes which are only more or less spherical. Kitaigorodsky showed, on the basis of a consideration of the packing of ellipsoids, that in these cases close packing is achieved in space groups containing as their principal symmetry elements glide planes and diad screw axes; this results in low symmetry space groups. Such are the rhombic groups $P2_12_12_1$, *Pbca*, *Pca*, the monoclinic $P2_1$, $P2_1/c$, the triclinic $\bar{1}$ and some others. The frequency of these groups can be judged from the following figures: the group $P2_1/c$ accounts for about 9% of crystals, and the group $P2_12_12_1$ for 4%.

Shubnikov Groups

The 230 Fedorov groups describe the symmetry of a periodic scalar field whose value at each point is determined by a single parameter. In order to describe the symmetry of a periodic vector field, which is determined by three components at each point, or of a tensor field, it is necessary to broaden the concept of symmetry, and this leads to an increase in the number of symmetry groups. An important generalisation of the concept of symmetry was suggested by Shubnikov, who introduced the idea of positive and negative or black-white figures besides the monochromatic figures considered in previous symmetry theory. The use of sign or colour in a figure (in the general case each point in space can have l 'signs', where l is a positive integer) leads to the introduction of the new concept of *antiequality*, the relationship between figures with different signs, and the new symmetrical transformation of *antisymmetry*, which changes the sign of a figure. The operation of antisymmetry, when transforming the various figures, changes a positive figure into a negative one and vice versa. The application of this operation to self-symmetric figures renders them neutral.

Latterly Soviet crystallographers have derived all point and space groups by using the operation of changing the sign of a figure. The system of generalised symmetry groups thus obtained has applications in the description of the symmetry of dipolar structures (magnetic structures, ferrites etc.), the symmetry of generalised projections of crystals, where the maxima may be positive or negative, and so forth. The generalised point groups together with the 32 ordinary ones contain 58 black-white magnetic groups. The generalised space groups contain in all 1651 groups, including 230 monochromatic and 230 grey Fedorov groups, and in addition 1191 black-white groups. The black-white groups describe the symmetries of periodic magnetic structures, and the grey groups that of paramagnetic crystals. The number of Bravais lattices also increases if we take antisymmetry into account. Apart from the 14 ordinary types, there are 22 black-white Bravais lattices.

5. The Symmetry of Atoms and the Types of Simplest Periodic Structures formed by them

Let us consider the simplest periodic structures arising when identical atoms are joined together. The type of structure is determined by the symmetry of the atoms and the properties of the bonds formed by them. With the homopolar bond we have to distinguish two cases: atoms with directed and saturated bonds (covalent bond) and atoms with undirected and unsaturated bonds (metallic bond).

Atoms with Directed and Saturated Bonds

As models for these atoms we can take little spheres with some of their radii produced to represent the lines of the bonds. Let us consider atoms forming (*a*) a single bond; (*b*) two bonds: linear (valence angle $\phi = 180°$) and orthogonal ($\phi = 90°$); (*c*) three bonds: plane ($\phi = 120°$) and pyramidal ($\phi = 90°$); (*d*) four tetrahedral bonds uniformly distributed in space ($\phi = 109° 28'$).

Atoms with a single bond are united into diatomic molecules, which on further interaction form a molecular crystal, for example the crystals of the halogens Cl_2, Br_2, I_2 (see Fig. 86*a* (171). Atoms with two bonds form chains which may be either infinite and unclosed (selenium and tellurium lattices) or closed in rings with a finite number of links (the molecules S_8) (see Fig. 86*b* and *c* (171)). Atoms with three plane bonds form infinite plane layers consisting of hexagonal loops (Fig. 81*a*). The unit cell of such a network is a rhombus comprising two particles with coordinates (0, 0) $(\frac{2}{3}, \frac{1}{3})$. The graphite lattice is composed of plane atomic networks as shown in Fig. 86*f* (171). In the case of three pyramidal bonds similar layers of hexagons are obtained, but here the atoms are arranged on two different levels. Such layers are found in the lattices of arsenic, antimony and bismuth (Fig. 86*d* (171)). Atoms with tetrahedral bonds form a three-dimensional spatial structure of the diamond type (Fig. 81*b*). The elementary cell of this structure is cubic, and it contains 8 atoms; the Bravais lattice is face-centred (Fig. 86*e* (171). The coordinates of the initial nodes of the two interpenetrating face-centred lattices are (000), $(\frac{1}{4}\,\frac{1}{4}\,\frac{1}{4})$.

The atomic network of graphite and the structure of diamond are examples of lattices of identical particles. These lattices cannot be formed by means of translations alone, without rotations and reflections.

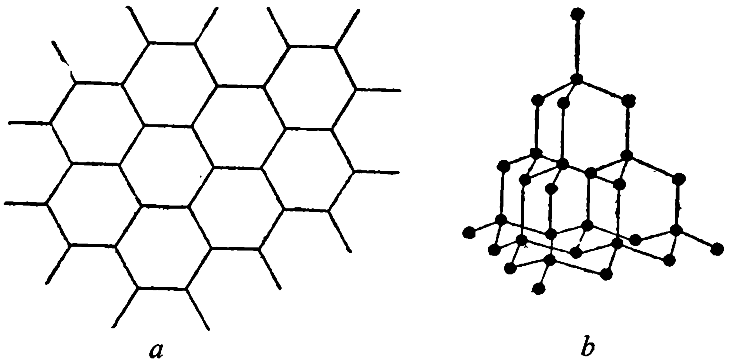


FIG. 81. (*a*) Atomic network formed by atoms with three coplanar bonds (graphite lattice); (*b*) space structure formed by atoms with tetrahedral bonds (diamond structure).

Structures formed by particles with saturated bonds are characterised by low coordination numbers. In crystals consisting of diatomic molecules the coordination number $k = 1$. In crystals composed of linear chains or rings $k = 2$. In stratified crystals like graphite or arsenic $k = 3$. In structures of the diamond type $k = 4$.

Spherical Particles with non-directed Bonds. Close packing of Spheres

Particles with undirected and unsaturated bonds may be represented to a first approximation by equal incompressible spheres of radius R, which attract each other. From the principle of minimum potential energy of interaction it follows that each particle tries to interact with the maximum possible number of other particles, and this leads to the formation of structures with maximum coordination numbers. Accordingly the *packing fraction*, that is, the ratio of the volume of the spheres to the total volume (including the voids between the spheres), will be larger than that of any other possible packing. Packings of spheres which satisfy these requirements will be called close packings. *One-dimensional close packing* is a linear periodic chain of spheres with period $2R$ and coordination number two. *Two-dimensional close packing* (plane spherical layer) is shown in Fig. 82*a*. This represents a two-dimensional, periodic structure with hexagonal symmetry and coordination number six. The plane unit cell of this structure is a right rhombus $AAAA$ with sides $2R$ (Fig. 82*b*). Between the spheres in the plane layer two families of trigonal (triangular) voids $BB\ldots$ and $CC\ldots$ are formed, across which pass two families of threefold axes of symmetry. Each of these families forms a similar plane hexagonal network with the same period $2R$. The point symmetry group for the positions B and C, $C_{3v}-3m$, is a subgroup of the symmetry of the position A, $C_{6v}-6mm$.

Three dimensional close packing is obtained by a compact combination of the

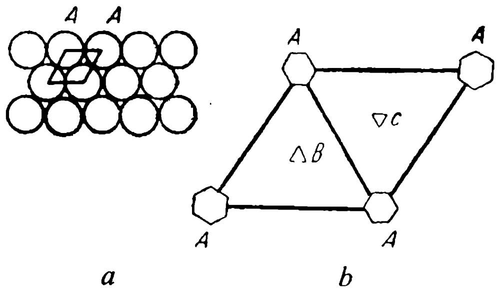


FIG. 82. (*a*) Close-packed layer of spheres; (*b*) unit cell.

close packed plane spherical layers. We take an initial plane layer, which we assume to be horizontal, and lay on it a sphere of a second layer. This sphere attains its most stable position if it is placed in one of the trigonal depressions. In this position, a sphere of the second layer is in contact with the greatest possible number, namely three, spheres of the first layer, forming with them a right tetrahedron. We continue adding spheres to the second layer in such a way that they in their turn also form a close packed plane layer. The spheres of the second layer will occupy one of the families of trigonal holes and will be in close contact both with spheres of the first layer and with each other. In the packing of two neighbouring layers the alternatives of placing in accordance with either of the families of holes *B* and *C* are equivalent, but starting with the third layer these two methods of adding each succeeding layer lead to different types of space packings. Any three-dimensional close packing can be obtained by the successive threading of dense spherical layers onto the families of axes of symmetry *A*, *B* and *C* drawn perpendicular to the plane of the initial layer. Such a system of close packing is uniquely characterised by a symbol of the type . . . *ABACABC* . . . , in which each letter relates the positions of spheres of the corresponding layer to the axes *A*, *B* or *C*. From the condition for close fitting of neighbouring layers it follows that letter symbols with two adjacent letters identical are impossible. In three dimensional close packing, the coordination number $k = 12$. Any close packing of spheres conserves the minimal point symmetry $C_{3v}-3m$.

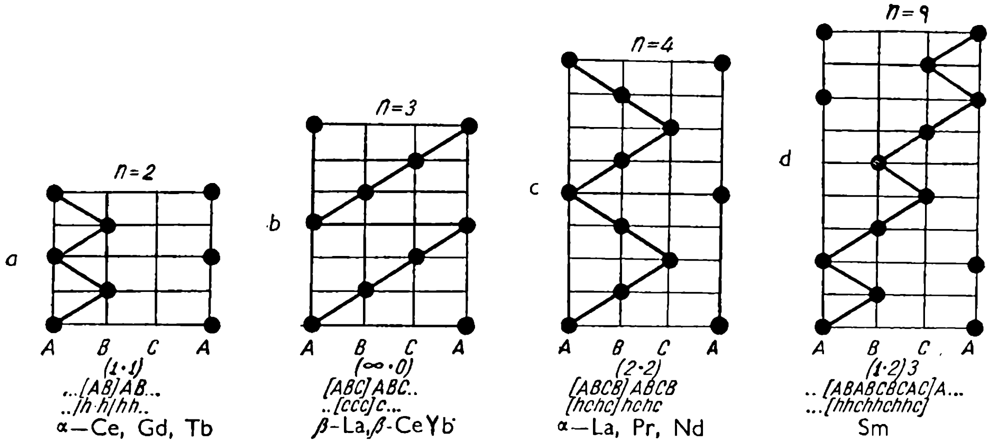


FIG. 83. Close-packing of spheres: (*a*) two-layer (hexagonal), (*b*) three-layer (cubic), (*c*) four-layer (hexagonal), (*d*) nine-layer (rhombohedral).

From the various kinds of three dimensional close packings we can distinguish the periodic packings, those which are described by periodic letter symbols. To indicate them it is sufficient to give the sequence of letters in a single period. The number *n* of letters in a period determines the stratification of the packing. With increasing stratification the number of close packings grows without limit. The packings of greatest physical importance are those with a low stratification. One of the simplest is the *two-layer* packing . . . *ABABAB* (Fig. 83*a*) (alternatively denoted *ACAC* or *BCBC* . . .), which is related to the hexagonal system. The structure of the hexagonal close packing Γ_h' is shown in Table 33.

The hexagonal close packing (h.c.p.), like the structures of graphite and diamond, cannot be obtained by means of translations alone. The unit cell of this packing, which has a ratio of axes $c/a = 2\sqrt{\frac{2}{3}} = 1{\cdot}633\ldots$, has two particles with coordinates: (000) $(\frac{1}{3}, \frac{2}{3}, \frac{1}{2})$, the space group is $D_{6h}^4 - C^6/mmc$.

The three-layer close packing is describable by only one letter symbol ... *ABCABCABC* ... (Fig. 83*b*). This is the only close packing related to the cubic system, and has the face-centred cubic structure Γ_c'' (see Table 33). The space group is $O_h^5 - Fm3m$. Unlike other close packings, the three-layer packing contains not one but 4 equivalent families of densely packed planes, parallel to the faces of the octahedron [111]. Cubic close packing is an example of a structure in which the positions of the particles coincide with the positions of the nodes in a Bravais lattice and provide an immediate model of the subgroup of parallel translations. Figs. 83*c* and *d* show the more complex 4- and 9-layer packings; the first is hexagonal and the second rhombohedral. These packings are encountered amongst the structures of the lanthanides (Ch. 6, §1).

In Fig. 80 a primitive rhombohedron with volume equal to $\frac{1}{4}$ that of the cubic cell and containing a single particle has been marked out. This rhombohedron can be dissected into an octahedron and two tetrahedra possessing identical triangular faces with sides equal to the shortest periods of the close packing, $2R$. Thus, space can be completely filled without any voids by a combination of octahedra with double the number of tetrahedra. The geometrical meaning of the octahedra and tetrahedra in close packed lattices is that the centres of these polyhedra correspond to the centres of the two types of greatest voids between the spheres; these are called octahedral or tetrahedral according to the type of coordination formed by the spheres which surround them. The number of octahedral voids is equal to the number of spheres in the packing; the number of tetrahedral voids is double that. Any system of close packing can be represented by a model composed of such polyhedra combined in a definite order.

In the formation of some compounds the voids are for one reason or another filled by atoms of a different chemical sort. The atomic structures of such compounds can nevertheless be conveniently described by the theory of close packing.

The symmetry of the spheres differ between the hexagonal and the cubic close packings because of the difference in the configurations caused by the different packings of neighbouring layers. The two ways of adding a layer of spheres give rise to two 'types of spheres': those with hexagonal symmetry (type *h*) and those with cubic symmetry (type *c*). If a spherical layer lies between layers with identical letter symbols, then it will contain spheres of type *h*. Horizontal planes of symmetry pass across them; the spheres do not have centres of symmetry; their point symmetry is $D_{3h} - 6m2$. If a layer of spheres lies between layers with different letter symbols, then it will contain spheres of type *c* with point symmetry $O_h - m3m$. These spheres possess centres of symmetry, but no horizontal planes of symmetry pass across them.

The hexagonal and cubic close packings can be denoted by the symbols .. *hhhhhh* .. and ... *ccc* .. Any other close packing is denoted by a mixed symbol containing the appropriate number of *h*'s and *c*'s in a definite order.

Numerical Symbol of a Spherical Close Packing

A symbol of the type *h*, unlike a symbol of the type *c*, causes a 180° change in the orientation of the layer; this is shown in the form of zigzags in Fig. 83*a*. Any periodic close packing is uniquely defined by the thicknesses of the successive layers with parallel and antiparallel orientations. We shall write this sequence as a *numerical symbol*

$\ldots p_1 \cdot p_2 \cdot p_3 \ldots p_j \cdot \cdot p_m \ldots$, where p_j is the number of elementary layers of spheres in the jth layer (packing). The point sign in the numerical symbol indicates the boundaries of the packings and is not a multiplication sign. The numbers p_j are positive integers; the number of layers in packing period m must be even. Cyclic permutation of the digits in the symbol, or writing the symbol in the reverse order, does not alter the type of packing. Numerical symbols differing in the sets or order of their numbers represent different packings.

The numerical symbol enables us to give a complete description of a spherical packing, using the following rules.

1. The stratification of the packing according to the shortest possible translation and the (equal) number of particles in a primitive unit cell is

$$n = \sum_{j=1}^{m} p_j. \qquad \ldots [5.6]$$

The stratification along the normal to the basal planes (i.e. along the axis c) is $n_c = n$ for hexagonal packings and $n_c = 3n$ for rhombohedral packings.

2. The dimensions of the unit cell of the packing in hexagonal axes are

$$c = n_c h, \quad a = \sqrt{\tfrac{3}{2}}\,h, \quad \frac{c}{a} = n_c\sqrt{\tfrac{2}{3}} = 0{\cdot}817\,n_c, \qquad \ldots [5.7]$$

where $h = 2R\sqrt{\frac{2}{3}}$ is the height of a tetrahedron with sides $2R$.

3. The type of Bravais lattice is determined by the difference between the total thicknesses with different orientations.

$$\Delta = \Sigma\, p_{2j+1} - \Sigma\, p_{2j} = n - 2\,\Sigma\, p_{2j}. \qquad \ldots [5.8]$$

With $\Delta = 3k$ the primitive lattice is hexagonal, and with $\Delta = 3k+1$ it is rhombohedral.

4. The coordinates of the particles in hexagonal axes are $\left(0, 0, \dfrac{p}{n_c}\right)$ $\left(\frac{2}{3}, \frac{1}{3}, \dfrac{q}{n_c}\right)$ $\left(\frac{1}{3}, \frac{2}{3}, \dfrac{r}{n_c}\right)$ and are determined by the displacements p, q, r of the particles along the axes of symmetry A, B and C. The numbers p, q, r (the numbers of the layers) can be found from graphs like the ones shown in Fig. 83, which can easily be drawn given the numerical symbol.

5. The symmetry elements of the packing are determined by the symmetry of the numerical symbol:

(*a*) if the packing symbol contains two identical half-periods with an odd number of members, then the packing has a two-way six-fold screw axis of symmetry 6_3 (space groups D_{6h}^4 and C_{6v}^4); in other cases the packings contain only three-fold axes of symmetry;

(*b*) if the packing symbol is mirror-symmetrical relative to the boundary between the layers (points), then there is a plane of symmetry passing between the corresponding layers (groups D_{6h}^4 and D_{3h}^1);

(*c*) if the packing symbol is mirror-symmetrical relative to any layer (digits), then there are centres of symmetry lying in the middle of the layer (groups D_{6h}^4, D_{3d}^3, D_{3d}^5).

Table 27

Classification of Spherical Close Packings

Symmetry		Hexagonal			Trigonal				Rhombohedral	
Space group		D^4_{6h}–$C6/mmc$		$C^4_{6v}C$–$6m^4_6$	D^3_{3d}–$C\bar{3}m$	D^1_{3h}–$C\bar{6}m$		C^1_{3v}	D^5_{3d}–$R\bar{3}m$	C^5_{3v}–$R3m$
Symmetry elements		×	×		× × ×				× × ×	
Stratification (Repetition of layers) (n_c for hexagonal packing, $\frac{1}{3}n_c$ for rhombohedral packing)	2	1	1							
	3								(1 · 2)3	
	4	2	2						(1 · 3)3	
	5				1 4				(2 · 3)3 (1 · 1 · 2 · 1)3	
	6	3	3			1 · 2	2 · 1		(1 · 5)3 (2 · 4)3 (1 · 1 · 3 · 1)3	
	7				2 · 5 4 · 1 · 1 · 1 3 · 1 · 2 · 1				(1 · 6)3 (3 · 4)3 (1 · 2 · 2 · 2)3 (1·1·1·1·1·1·2)3	(1 · 2 · 1 · 3 ·)3
	8	4 1·2·1	4 1·2 1		7 1 · 2 · 3 · 2	1 · 3 1·1·2	3 · 1 2·1·1		(3 · 5)3 (2 ·)3 (1 · 1 · 1 · 5)3 (1 · 2 · 1 · 4)3 (2·1·1·1 · 2 · 1)3 (3·1·1·1 · 1 · 1)3	(1 · 1 · 2 · 4)3 (1 · 2 · 2 · 4)3

The above-mentioned symmetry elements uniquely define the appropriate space group out of the total of seven single-axis space groups D_{6h}^4, D_{3d}^3, D_{3d}^5, C_{6v}^4, D_{3h}^4, C_{3v}^4, C_{3v}^5 which, according to Belov, are possible for close packings. Each of these groups contains infinitely many close packings. Table 27 gives the classification of packings up to and including the 8-layer type; the crosses and two parallel lines show the centres of symmetry and planes of symmetry.

The stratifications of rhombohedral packings are trebled; this corresponds to the description of these packings in terms of hexagonal axes. The table does not include the eighth possible space group $O_h^5 - Fm3m$, which is associated with cubic close packing. This packing does not contain spheres of the type h, successive layers in it are parallel and it is a limiting case characterised by the numerical symbol $\infty \cdot 0$. Packings with the symmetry C_{3v}^1 are first encountered among the 9-layer ones: $1 \cdot 1 \cdot 2 \cdot 5$; $1 \cdot 2 \cdot 2 \cdot 4$; $1 \cdot 1 \cdot 1 \cdot 2 \cdot 1 \cdot 3$. Packings with the symmetry C_{6v}^4 appear when we reach $n = 12$, and are $1 \cdot 2 \cdot 3 \cdot 1 \cdot 2 \cdot 3$.

A number of crystalline substances crystallise in close packed form. Of the simple substances, metals in many cases crystallise in cubic ($\infty \cdot 0$) or hexagonal ($1 \cdot 1$) close packing. Neodymium and samarium crystallise in more complex packings: $4-x$ ($2 \cdot 2$) and 9-layer ($1 \cdot 2$) 3 respectively. Of the binary compounds, NaCl and ZnS (sphalerite) (Fig. 96) and many substances isomorphic to them crystallise in cubic close packing with filling of the octahedral and tetrahedral voids respectively. The other modification of ZnS (würtzite) has a structure based on hexagonal close packing. In Al_4C_3 the aluminium atoms form a 9-layer close packing ($1 \cdot 2$) 3; in $CaSi_2$ the silicon atoms form the 12-layer packing ($1 \cdot 3$) 3. The exceptional variety of close packings is shown in the polymorphic and polytypal modifications of silicon carbide SiC, for which hexagonal ($2 \cdot 2$), ($3 \cdot 3$), ($4 \cdot 4$) and rhombohedral ($2 \cdot 3$) 3, ($2 \cdot 3 \cdot 3 \cdot 3$) 3, ($2 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$) 3 structures etc. have all been established. The theory of close packing has made it possible to decipher a number of complex structures belonging to minerals, metallic compounds and the superperiodic structures of silicon carbide, zinc sulphide etc.

Coordination Number and Packing Fraction

The packing fraction q of a structure is equal to the ratio between the volume of the particles from which the crystal is formed and the total volume of the crystal:

$$q = \frac{\sum_i Z_i v_i}{v_a}, \qquad \ldots [5.9]$$

where Z_i is the number of particles of type i in unit cell, v_i is the volume of a single particle, v_a that of a unit cell, and the summation extends over all particles in the cell. For spherical packings formed from identical spheres, the formula simplifies to

$$q = \frac{4\pi Z R^3}{3 v_a}. \qquad \ldots [5.10]$$

Table 28 gives the values of q for a number of structures made up of equal spheres with different values of the coordination number k.

As k decreases, so does the packing fraction. The packing fraction attains its maximum value ($q = 74\%$) in the densest packings, for which about $\frac{1}{4}$ of the total volume is accounted for by the voids between the spheres. In the simple cubic structure Γ_c the volume of the voids constitutes almost half the volume of the crystal. In the diamond structure the volume of the spheres is only $\frac{1}{3}$ of the total volume of the crystal. From the point of view of compactness this structure is extremely crumbly. In the b.c.c. structure Γ_c' q is nearer to its value $q = q_{\max}$ in the closest packings than to its value in the simple cubic structure, although the coordination number of the structure Γ_c' is nearer to the coordination number of the structure Γ_c than to that of Γ_c''. This is connected with the fact that in the structure Γ_c' the second coordination sphere is close to the first, the radii of these spheres ($r_{200}:r_{111} = \sqrt{4}:\sqrt{3}$) differing by only 15%. Although the interatomic interaction diminishes rapidly with increasing distance and the main part of the energy of a homopolar crystal is due to the interaction of nearest neighbours situated on the first coordination sphere, in a structure of the type Γ_c' there is appreciable energy due to interaction with particles of the second coordination sphere. So in this structure the effective coordination number $k = k_1 + k_2 = 8 + (6)$ has a value greater than 8. A high packing fraction $q \approx 70\%$ is found in the body-centred tetragonal structure Γ'' with a ratio of axes $c/a = \sqrt{\frac{2}{3}} = 0{\cdot}817$. The structure Γ_t'' is obtained by compressing the structure Γ_c' along one of the edges of the cube by 18%; this leads to an increase in the coordination number calculated from the first coordination sphere to $8+2 = 10$. In the tetragonal structure Γ_t'' the two mutually orthogonal families of planes (110) and ($1\bar{1}0$) are planes of the close packed type. From the condition for close packing of spheres, $a_t\sqrt{2} = c_t\sqrt{3}$ we obtain the above-mentioned ratio of axes for a tetragonal structure with coordination number $k = k_1 = 10$. This is the structure found in protactinium.

Table 28

Packing Fractions q for Lattices made up of Equal Spheres

Lattice type	Symbol	k	q	Numerical value of q %
close packing of spheres	Γ_c'', Γ_h' and others	12	$\frac{\sqrt{2}\pi}{6}$	74·04
tetragonal $\frac{c}{a} = \sqrt{\frac{2}{3}}$; $z = 2$	Γ''	10	$\frac{2\pi}{9}$	69·8
body-centred cubic	Γ_c'	8 ✓	$\frac{\sqrt{3}\pi}{8}$	68·1
simple cubic	Γ_c	6	$\frac{\pi}{6}$	52·3
diamond type	Γ_c^a	4	$\frac{\sqrt{3}\pi}{16}$	34

Problems

1. How many particles are there in the unit cell of (*a*) a body-centred; (*b*) a face-centred lattice?
2. Calculate the packing fraction for close packing of equal circles on a plane.

Chapter 6

THE STRUCTURE OF CRYSTALS: CRYSTAL CHEMISTRY

1. Crystal Chemistry of the Periodic System of the Elements

Consider qualitatively the simplest structural and chemical characteristics of an element, the coordination number k and basic valency v. From these crystal-chemical indications two regions can be distinguished in the periodic table (Table 29): on the left-hand side and in the central parts of the table including the lanthanides and actinides, the relationship $k > v$ is characteristic of elements with the metallic type of bonding. On the right hand side of the table the relation $k = v$ is characteristic of elements which tend to form covalent bonds. Some elements (carbon, nitrogen, oxygen and sulphur) in the early periods of this part of the table form structures with lower multiplicity $k < v$. Atoms of these elements are capable of forming multiple bonds.

Elements of the A Subgroups

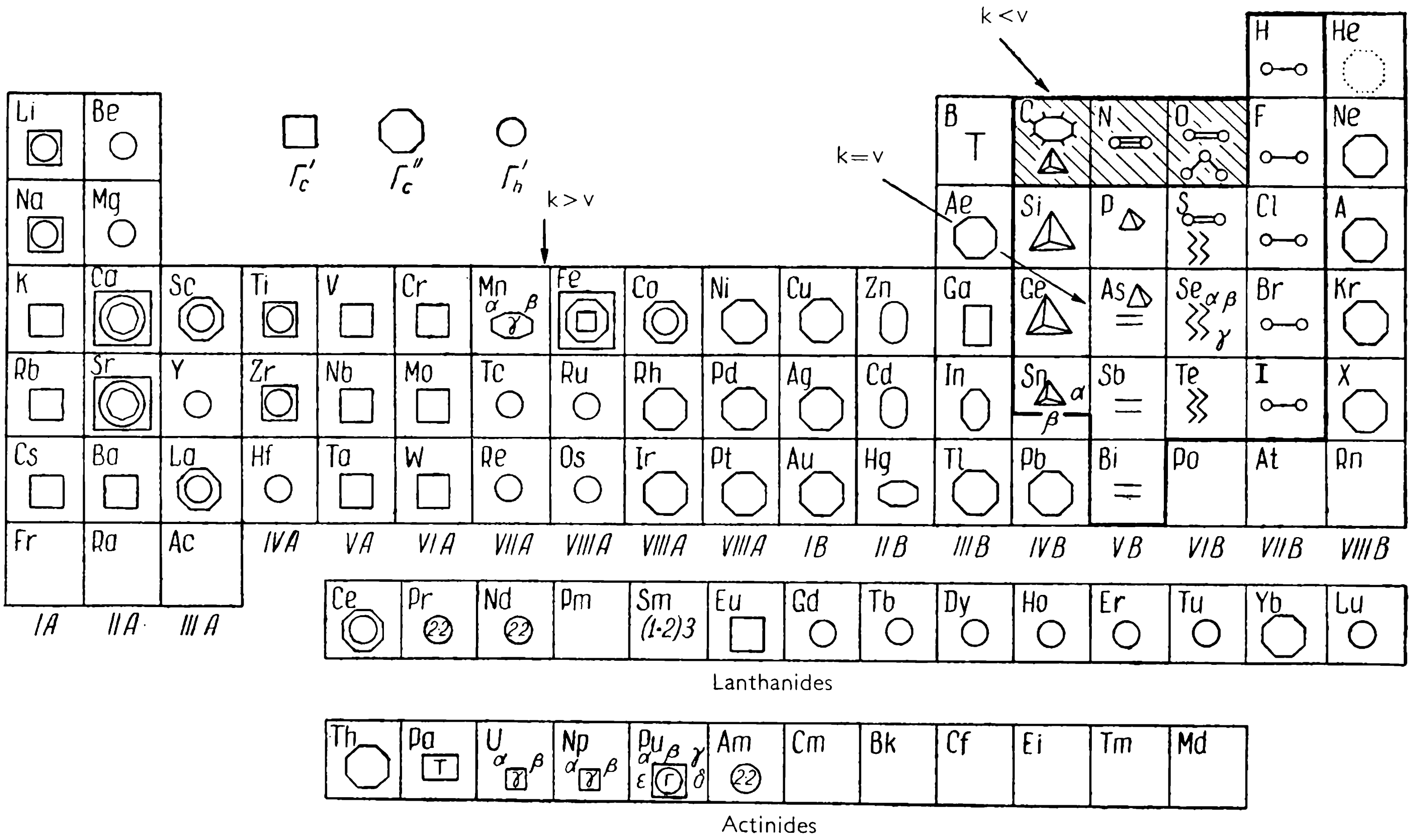
Among the elements of the A subgroups three types of structure predominate: cubic close packing (f.c.c.—indicated in Table 29 by an octagon) hexagonal close packing (h.c.p.—a circle) and body centred cubic (b.c.c.—a square). The difference in energy between these structural types is small so, among the metals, polymorphism is widespread. In cases where there is polymorphism, the innermost symbol (Table 29) refers to the lowest temperature modification.

The **alkali metals** are all b.c.c., but recently, low temperature modifications of lithium and sodium have been found which are h.c.p. and plastic deformation of lithium at low temperatures has been shown to lead to the appearance of a new f.c.c. modification. Calcium and strontium exist in three modifications.

The **transition metals** show the following regularities: when the filling of the d-electron shells passes the half way mark, modifications arise with close packed structures; if the d-shells are half-filled (chromium sub-group) or somewhat less full (vanadium and titanium sub-groups) the b.c.c. structure arises. We have to exclude from these the b.c.c. structure of low and high temperature modification of iron and also the structure of the polymorphic modification of manganese, where, in the cubic unit cell there are 58 atoms occupying 4 structurally different positions. The structure of α-Mn is analogous to that of the intermetallic compound $Mg_{17}Al_{12}$. β-Mn also crystallizes in the cubic system; the unit cell contains 20 atoms, occupying 2 struc-

Table 29

Crystal Chemistry of the Periodic System


Γ_c'
Γ_c''
Γ_h'
k > v
k = v
k < v
H He
Li Be B C N O F Ne
Na Mg Ae Si P S Cl A
K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn Ga Ge As Se Br Kr
Rb Sr Y Zr Nb Mo Tc Ru Rh Pd Ag Cd In Sn Sb Te I X
Cs Ba La Hf Ta W Re Os Ir Pt Au Hg Tl Pb Bi Po At Rn
Fr Ra Ac
α β γ
IA IIA IIIA IVA VA VIA VIIA VIIIA VIIIA VIIIA IB IIB IIIB IVB VB VIB VIIB VIIIB
Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tu Yb Lu
22
(1·2)3
Lanthanides
Th Pa U Np Pu Am Cm Bk Cf Ei Tm Md
T
ε δ
Actinides

turally different positions so that the structure is analogous to that of the compound Ag_3Al. β-Mn is interesting in that its structure has no centres of symmetry while all the other metals, and, with rare exceptions, almost all intermetallic compounds have centro-symmetric structures. γ-Mn has a structure of the f.c.c. type, but somewhat squashed as a result of 6% contraction along one of the cube edges giving a tetragonal face centred structure with $c/a = 0{\cdot}934$ and a similar structure to that of the ordered modification of AuCu.

The **lanthanides** crystallize in forms corresponding to close packing of spheres. Here, for the first time amongst simple substances, there appears a form of packing with a higher degree of complexity than the two-layer h.c.p. (1.1) and the three layer f.c.c. (∞ . 0). This is a four-layer hexagonal packing (2.2) (Fig. 83*c*) which is the crystal structure of α-lanthanum and also of praesodymium and neodymium. Lanthanum is dimorphic and transforms above 300°C into the β-modification which has the f.c.c. structure. Samarium crystallizes in nine-layered rhombohedral packing (1.2)3 (Fig. 83*d*). There are interesting data on the polymorphic transition of cerium under the influence of pressure: at 12,430 atmospheres cerium acquires a much smaller volume and it has been established by X-rays that this transition, although accompanied by a considerable change of volume $\left(\frac{\Delta v}{v} = 10\%\right)$ is not associated with a change of structural type: both above and below the transition point the structure of cerium is the same—f.c.c. It seems that the transition corresponds to the transfer of one of the 4*f* electrons to the 5*d* valence level, i.e. it is connected with the excitation of a new valence state of cerium. From spectroscopic data the energy of the transition $4f \to 5d$ is small ($\Delta E \sim 0{\cdot}04$ eV) and the transition can be stimulated by external pressure which causes an overlap of neighbouring electron levels. It is known that cerium has, in addition to trivalent compounds, e.g. Ce_2O_3 analogous to those of the other neighbouring lanthanides, also tetravalent compounds (CeO_2). A reduction of the atomic radius of cerium from 1·82 to 1·71 Å produced by external pressure corresponds to a difference of the ionic radii of 3- and 4-valent cerium (Ce^{III}—1·02 Å and Ce^{IV}—0·88 Å).

The **actinides** where, beginning with neptunium there is an overlap of the 5*f*, 6*d*, 7*s* and 7*p* valence levels have extremely complex structures. The number of polymorphic modifications increases rapidly as we pass from thorium to plutonium, reaching in plutonium the record number of six. Fig. 84 gives characteristic structures of the actinides: thorium has the f.c.c. structure, protactinium has a distorted tetragonal body centred structure with axial ratio $c/a = 0{\cdot}825$ and coordination number $k = 10$. As can be seen from the diagram the coordination polyhedron of protactinium is obtained by removing two atoms (9 and 11) from the coordination polyhedron of thorium which has a cubic octohedral shape. The high temperature modifications of uranium, neptunium, and plutonium have the b.c.c. structure. The low temperature modification of α-uranium belongs to the rhombic system with four atoms in the unit cell. The structure is an isotype of the intermetallic compound CdAu and represents a very heavily compressed hexagonal close packing. As a result of this compression the coordination number is reduced to four. The relation to the coordination polyhedron of hexagonal packing is shown in Fig. 84. The configuration of atoms in α-uranium corresponds to a trigonal bipyramid from which one atom has been removed. β-uranium crystallizes in the tetragonal system with 30 atoms in the unit cell, occupying 7 different structural positions. The structure is an isotype of the σ phase encountered amongst the compounds of transition metals (FeCr, FeV etc.).

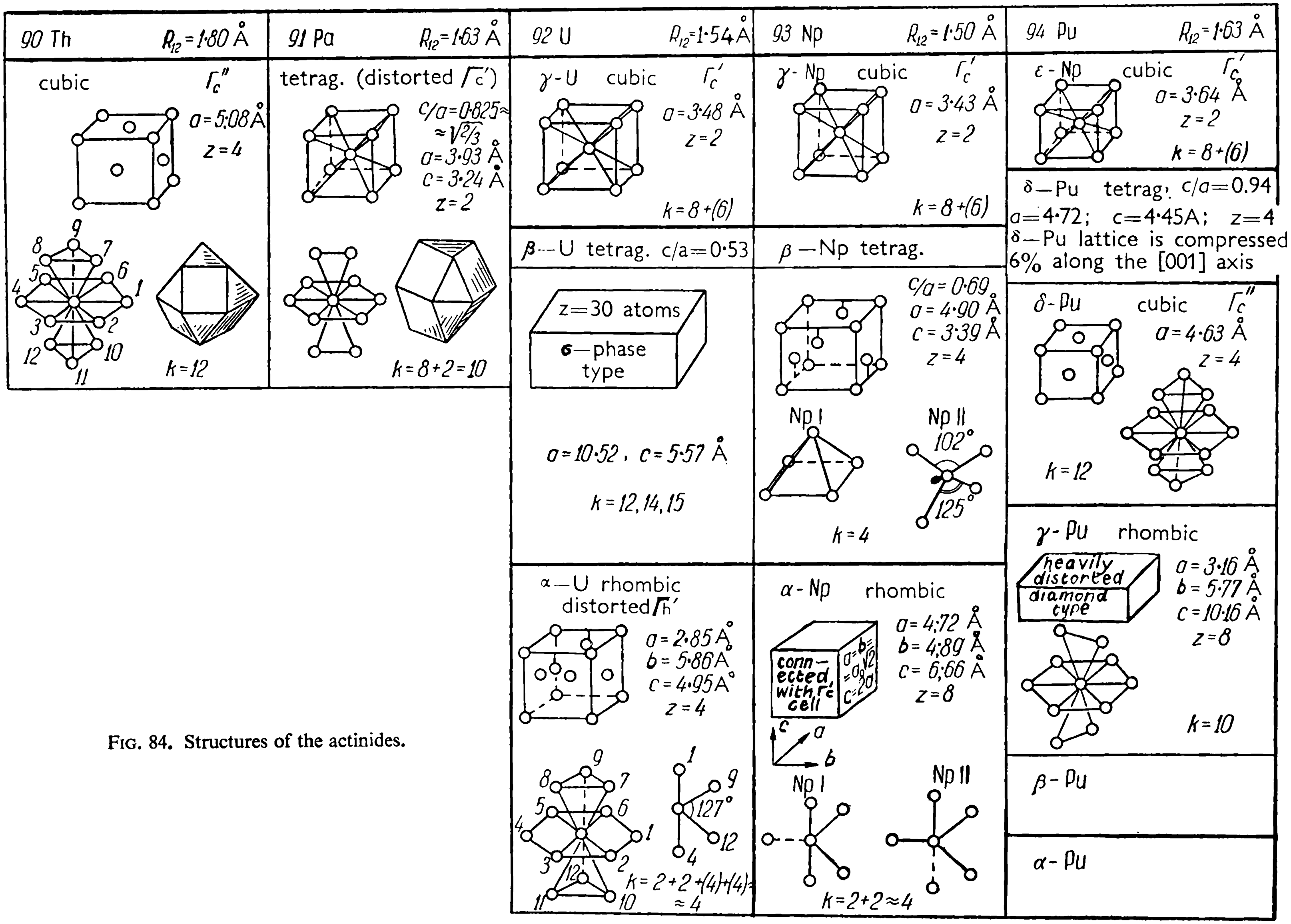


FIG. 84. Structures of the actinides.

α-neptunium has a rhombic unit cell containing 8 atoms occupying two structurally different positions. The configurations around the neptunium atoms correspond to a trigonal bipyramid from which either the side or the top atom has been removed (Fig. 84). In the former case the configuration is the same as in α-uranium. β-neptunium crystallizes in the tetragonal system and the unit cell contains 4 atoms occupying two structurally different positions. The coordination number is 4 as in α-neptunium but the configurations are different: a tetragonal pyramid and a distorted tetrahedron.

γ-plutonium has a rhombic cell containing 8 atoms. The coordination number $k = 10$. The coordination polyhedron is obtained from the polyhedron of close packing by taking away 2 atoms (Fig. 84). δ-plutonium has the f.c.c. structure. The strange thing about this modification is that, in all the regions where it exists, it has a negative coefficient of thermal expansion $\alpha_{\text{mean}} = -8{\cdot}6 \times 10^{-6}/°\text{K}$. This, it appears, is connected with a thermally produced perturbation of the highest valence state of the atoms of plutonium. The atomic radius of the high temperature modification, ε-plutonium is less than the atomic radius of the lower temperature modification, δ-plutonium, even though the change in coordination number has been put right. The intermediate modification δ'-Pu has the structure of δ-Pu but flattened by 6% along one of the edges of the cube. This structure is analogous to that of γ-manganese. Americium has a hexagonal four-layered packing (2.2) analogous to the structure of praesodymium and neodymium.

Because of the link with polymorphism the structure of thin metallic films has become especially interesting following on the recent investigations on low temperature polymorphism which arises under the influence of plastic deformation (Khotkevich). Electron diffraction studies of the structure of thin crystalline layers often reveal new structural types along with the usual modifications revealed by X-rays. Similarly, electron diffraction investigation of thin metallic foils show the existence of unusual structural types. Thus in the case of f.c.c. metals it was discovered some time ago (Quarrell, 1937) that there were also structures of the h.c.p. type. Recently Aggarwal and Goswani (1957) discovered that molybdenum, which usually has the b.c.c. structure, also had an f.c.c. modification. Uranium in the form of a thin film was found (Chatterjee 1958) to have a β-modification at room temperature, whereas this modification normally exists only in the temperature range from 660°-760° C.

The problem of the structure of thin metallic films is of great importance for the explanation of the nature of their peculiar properties. Thus, for example, thin films of Be and Bi exhibit superconductivity, while the common macroscopic modifications are not super conducting. Gallium in thin film form has transition temperature T_k so different from that of a macroscopic specimen of gallium that it is reasonable to assume that there is an essential difference of structure. In thin films of Bi and Ga condensed at 4·2° K there are changes in the lattice parameter (Buckel 1954).

Atomic Radii of the Metals

The majority of metals have simple symmetrical structures in which the atoms are structurally equivalent. This makes it possible to define an important structural characteristic of atoms, that is, the crystal-chemical atomic radius which is equal to half the shortest interatomic distance and is a measure of the strength of binding. The atomic radius depends on the coordination number.

Values of the atomic radii of metals M 12 with coordination number $k = 12$, and values of the corrections to the coordination number are given in Table 37 (see pages 192-3). In Fig. 85, curves are reproduced which show the dependence of the atomic radius M 12 on the number of electrons in the valence shell for elements in various periods of the periodic table. As the atomic number and the number of valence electrons is increased at the beginning of each period there is a marked reduction in

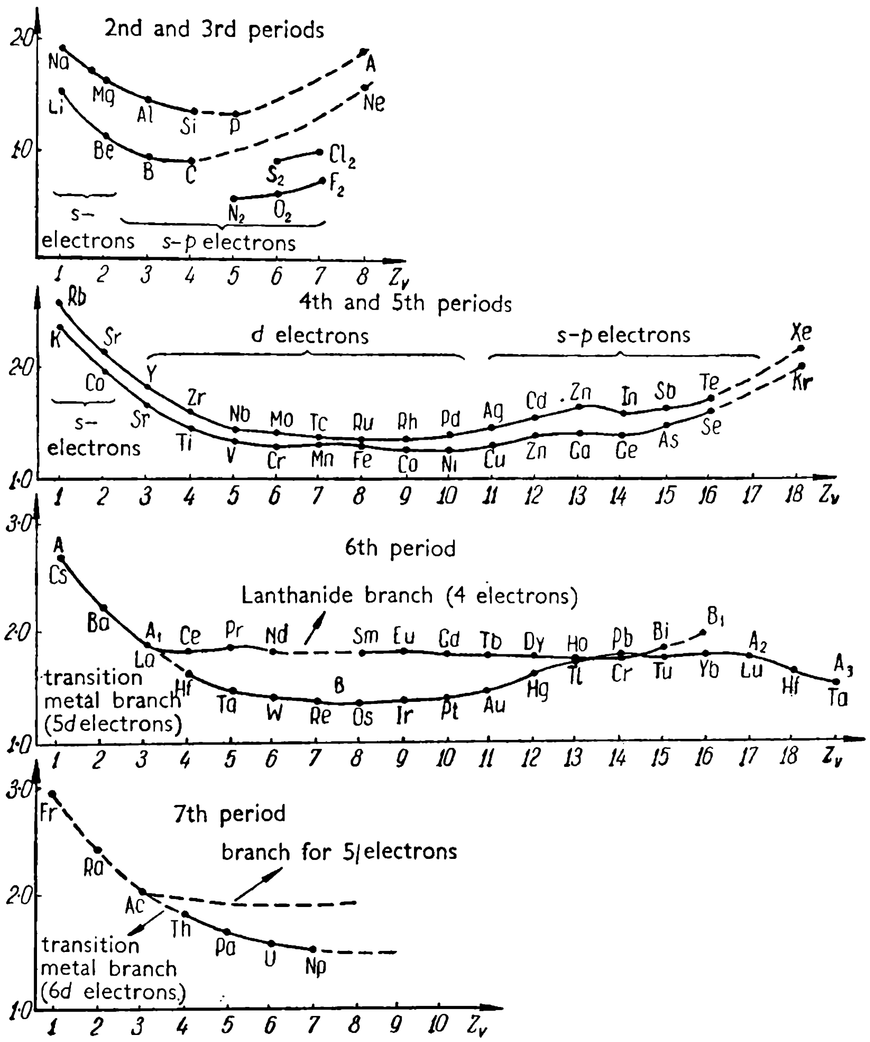


FIG. 85. Curves of atomic radii.

the atomic radius and at the same time an increase in the melting temperature. As the valence shells fill up, the curves of atomic radius pass through a minimum and then rise towards the end of the period. The transition from the second to the fifth period is accompanied by a movement of the corresponding curves parallel to the axis of ordinates; this is connected with an increase in the dimensions of the atoms.

In the **sixth period** the behaviour of this curve is interesting for it forms two branches. The basic curve $AA_1A_2A_3$ has a plateau which corresponds to the rare earth elements. The gradient of this plateau is small and thus shows that the 4*f* electrons make practically no contribution to binding. After the 4*f* shell is filled, the curve of atomic radii begins to fall rapidly (branch A_2A_3) and this is connected with the increase in the number of 5*d* electrons. If we exclude the rare earths from the curve of atomic radii, that is we move branch A_2A_3 fourteen elements to the left, then we obtain curve AA_1BB_1 which is very similar to the curve of the fifth period. The difference is that the sixth period curve has a deeper minimum in the region where it is almost coincident with the fifth period curve. This lowering of the values of the atomic radii of the elements following the rare earths is called the *lanthanide contraction*. As a result of it, the lattice parameter of gold ($a_{25^\circ} = 4{\cdot}0704$ kX) is less than that of silver ($a_{25^\circ} = 4{\cdot}0779$ kX).

The initial parts of the curve of the atomic radii of the elements of the **seventh period** are now known. By analogy with the sixth period we can draw the plateau which would correspond to the 5*f* electrons. However the experimental values of the atomic radii of the heavy elements from thorium to neptunium clearly lie on a lower branch, which corresponds to transition metals with 6*d* electrons.

Elements of the B-Subgroups

The elements grouped on the right-hand side of the periodic table have a much greater variety of structural types. As well as coordination structures we meet here, layer, chain and island (molecular) structures. Some of the elements have close-packed structures with large or small distortions. The complications of the pattern of structures in this region of the periodic table is a result of rapid change in the character of the binding, and of the transition from metallic unsaturated to covalent saturated binding, as we move towards the right hand side of the table. On the basis of structural data the elements of the B-subgroups can be divided into two parts: B_1 mainly metallic, B_2 mainly covalent. The line of division between them in general passes between the III B and IV B subgroups.

The characteristic features of the structures of the B_2 elements are determined in the first place by the symmetry of directed covalent bonds which are formed by atoms of corresponding elements. These structures have small coordination numbers which are determined by the octet rule

$$k = v = 8-n. \qquad \ldots [6.1]$$

where n is the number of the subgroup of the B_2 element. The number 8 here is equal to the electron capacity of the *s-p* levels. Crystals of the **inert elements** have the f.c.c. structure, which agrees with the spherical symmetry of the atoms of these elements. It follows from formula [*1*] that for the elements of group VIII, $k = v = 0$. Since there are only van der Waals' forces between the atoms of the inert elements, the crystals of these elements can be considered from the point of view of crystal chemistry as molecular crystals in which the molecules are the chemically inactive atoms. The rhombic crystals of the **halides** are constructed from diatomic molecules ($k = 1$) and are typically molecular (island). Thus in the chlorine crystal the spacing of atoms within each molecule is 65% less than the shortest distance between the atoms of the neighbouring molecules. These dumbbell molecules are distributed according to a face centred law (Fig. 86*a*), each molecule having 12 neighbours; the molecular packing is dense. Topologically, crystals of the halides and the inert elements all belong to the same type and the lowering of the symmetry of the crystals of the halides from cubic to rhombic is a result of lowering the symmetry of the structural elements arising from the transition from sphere to dumbbell.

The atoms of nitrogen and oxygen have a tendency to multiple bonding. Thus oxygen and nitrogen form stable diatomic molecules with double and triple bonds respectively. Oxygen also forms the less stable tri-atomic ozone molecule. In the solid state the molecules of these elements are preserved intact and the respective crystals are of the molecular island type. The atoms of carbon also have the possibility of forming multiple bonds. Since it is impossible from quantum chemical considerations to form fourfold bonds, the barely stable molecules of C_2 which have been shown to exist spectroscopically are unknown in the solid state. In the solid state the carbon atoms form three co-planar multiple bonds and this leads to the development of the layered structure of graphite (Fig. 86*f*). Structural data show that in the periodic table there is a small region of elements which includes carbon, nitrogen and oxygen which in the solid state form structures with lowered coordination $k<v$.

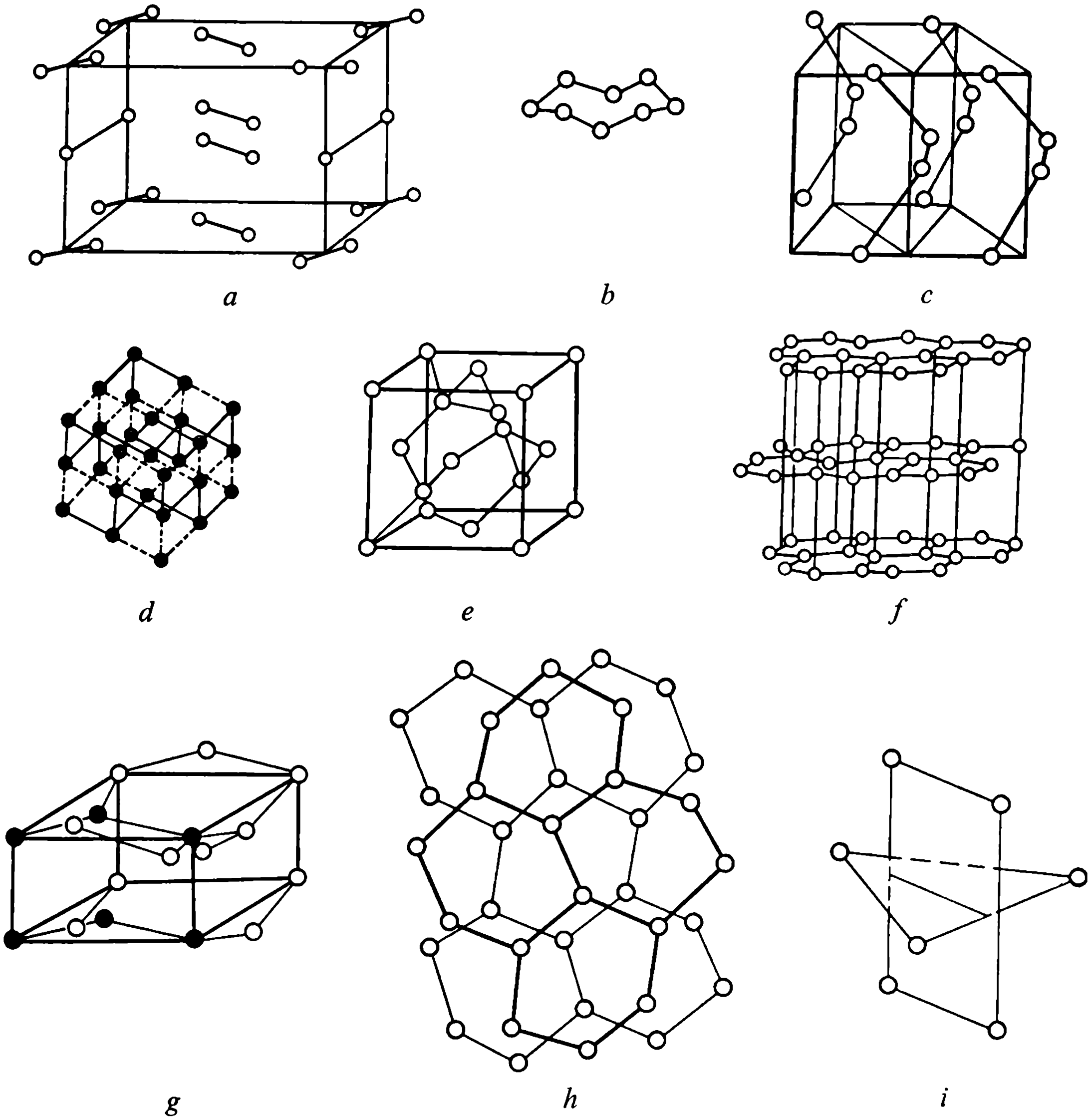


FIG. 86. Structures: (*a*) Cl_2 crystal, (*b*) S_8 molecule, (*c*) tellurium, (*d*) arsenic, (*e*) diamond, (*f*) graphite, (*g*) white tin, (*h*) gallium, (*i*) coordination polyhedron of gallium.

In the gaseous state, sulphur is known to form diatomic molecules of S_2 with double bonds. However, in the solid state the molecules S_2 do not survive. For the atoms of the elements of the VI B subgroup beginning with sulphur the formation of two single bonds is more favourable, the bonds being inclined at an angle to each other. This leads to the formation of zig-zag chains by the atoms of the elements mentioned. The chains can be closed as in the S_8 molecules (Fig. 86*b*) in the rhombic modification of sulphur and the Se_8 molecules in the monoclinic modifications of α and β selenium or not closed as in the plate-like modification of sulphur and the hexagonal structures of γ-Se and Te (Fig. 86*c*). The unclosed chain structures are denoted by the symbols S_∞, Se_∞, Te_∞. The valence angles in the $\alpha-$, $\beta-$ and $\gamma-$ modifications of selenium are $\phi \simeq 105°$. The valence angle in the tellurium structure is $\phi = 102°$.

Atoms of elements of the **fifth group** beginning with phosphorus form three bonds directed along the edges of a triangular pyramid. This leads to the formation of tetrahedral molecules P_4 and layer type crystals $As_{2\infty}$, $Sb_{2\infty}$, $Bi_{2\infty}$, which are built up of double atomic layers (Fig. 86*d*). Table 30 gives structural characteristics of elements of the V B subgroup.[†]

Table 30

Structural Characteristics of Elements of the VB Subgroup

Structural characteristics	As	Sb	Bi
Valence angles	73 and 97°	79 and 96°	80 and 93°
Ratio of short to long interatomic spacings	0·8	0·85	0·9

The atoms of the elements of **group IV** form four bonds directed along the apexes of a tetrahedron. This leads to the formation of three-fold coordination structures of the diamond type $C_{3\infty}$, $Si_{3\infty}$, $Ge_{3\infty}$ and $Sn_{3\infty}$ (Fig. 86*e*). As the atomic number of the atom is increased, the vacant energy band narrows leading to the development of metallic binding. This transition can be seen by the example of tin and lead. The structure of tetragonal white tin (β-tin) can be considered as a highly distorted diamond structure in which the coordination number has increased from 4 to 6. Lead is a typical metallic structure of the f.c.c. type. In group VI the structures of α- and β-polonium can also be considered as approximating to metallic: α-polonium is a simple cubic lattice and β-polonium is rhombohedral, being a slightly distorted cubic. Both have coordination number 6, i.e. considerably greater than in the chain structures of selenium and tellurium.

The **elements B_1** occupy an intermediate position between the metals (elements of subgroup A) and the elements B_2 with the covalent type of bonding. With the exception of boron and gallium these elements of group B_1 have simple or slightly distorted metallic structures. The elements of the copper subgroup are all f.c.c. Zinc and cadmium have structures of the h.c.p. type extended by 15-18% as a result of which the coordination number falls to 6+6. Mercury has a simple rhombohedral structure formed as a result of compression of an f.c.c. lattice by 50% along the body diagonal. The coordination number in the mercury crystal also falls to 6 nearest and 6 second nearest neighbours.

[†] As the atomic number of the element increases, the valence angle tends to the value $\phi = 90°$ which is characteristic of hybridisation. This can be explained by the reduction of the role of the *s* state in the bond as a result of the completion of the pairs of s^2 electrons (see page 82 and the footnote on page 119).

We meet with great variety and complexity of structure in the elements of **group III**. Aluminium and dimorphic thallium are typical metallic structures. Indium has a tetragonal structure which is produced by extending an f.c.c. lattice by 8% along one of the cube edges; the coordination number thereby falls to 4+8. In indium, nearest neighbours form a plane configuration in the form of a square, while in the neighbouring elements, zinc and cadmium, nearest neighbours also form a plane configuration but in the form of a hexagon. The tendency to the formation of plane configurations in crystals of these elements appears to be connected with the influence of the electrons of the *d*-subgroups.

Investigation of the structure of boron is a matter of considerable difficulty since it is not easy to obtain boron pure and in single crystals because of its high chemical activity and its high melting point. The needle-shaped crystals of boron have a tetragonal structure; the unit cell contains 50 atoms. In the structure, one can distinguish four groups of twelve boron atoms forming almost regular icosahedra arranged according to the tetrahedral law ($\frac{1}{4}, \frac{1}{4}, \frac{1}{4}$; $\frac{3}{4}, \frac{3}{4}, \frac{1}{4}$; $\frac{1}{4}, \frac{3}{4}, \frac{3}{4}$; $\frac{3}{4}, \frac{1}{4}, \frac{3}{4}$). Two other boron atoms are located on the apexes and in the centre of the cell. In the unit cell of the plate-shaped crystals there are about 200 atoms.

Gallium has a rhombic structure, the unit-cell containing 8 atoms situated in two parallel planes (Fig. 86*h*). In each plane the atoms form a net of irregular hexagons in which one pair of sides are shorter than the others. Of every three neighbours, one atom is the nearer and we can distinguish two pairs of neighbouring atoms from neighbouring planes. The total coordination number is 7. The coordination polyhedron has an irregular shape and can be thought of as the intersection of a rectangle with a triangle perpendicular to it (Fig. 86*i*).

2. Structures of Solid Solutions

Ways of distributing Atoms in Structures of Substances with Complex Compositions

In single phase materials with complex chemical composition there are three ways of placing the atoms. These we shall consider by the example of binary substances: (*a*) substitutional (A, A′) (*b*) interstitial (A, B), and (*c*) the formation of a new crystalline structure A_nB_m. Substances falling into the first two cases and for which the phase composition is variable are usually called solid solutions. In the third case we speak of the formation of a chemical compound or of a new crystalline phase if that phase forms solid solutions $(A, A')_n$ $(B, B')_m$. Phases of variable composition which do not obey the law of simple multiple proportions are often formed, not only in the gaseous and liquid, but also in the solid state. The majority of phases in metallic alloys, many compounds of the periodic metals with elements of the B subgroups and others, belong to phases with variable composition.

Whether substitutional phases or interstitial phases are formed depends on the chemical nature of the interacting atoms. For the substitution of any atom in the structure of another atom to be possible, it is necessary that the chemical properties of these atoms should be similar. The probability of forming substitutional solid solutions is greatest when the interacting atoms belong to the same subgroup of the periodic table, A, A′ or B, B′, and also when the interacting atoms are metals M and M′ from different subgroups. The latter case arises because of the 'impersonal' nature of metallic bonds which are the result of collectivisation of the valence electrons. These three cases for binary systems can be written in the form (A, A′), (B, B′) and (M, M′).

The formation of interstitial structures on the other hand requires a considerable difference in the properties of the interacting atoms A and B and of these, the most important is a difference in the atomic volumes. This is because the radii of the empty volumes between atoms in any structure are usually less than the radii of the atoms of that structure. Thus the atomic volumes of atoms going into interstitial positions must be less than the volumes of the atoms of the basic carcase of the lattice. This requirement for atomic volumes is necessary but not sufficient, since putting atoms into the structure is accompanied by chemical interactions and it is impossible to consider this purely geometrically like filling free volumes with spheres of suitable radius. In some complex compounds which have a crumbly structure there are quite large voids and the volume condition is not a factor limiting the possibility of inserting additional atoms or atomic groups (water molecules, ammonia molecules, and so on).

Substitutional Solid Solutions

Experimental results show that elements belonging to the same subgroup have a tendency to form substitutional solid solutions both in the case of metallic binding and in the case of covalent binding and in many favourable binary systems a continuous series of solid solutions is formed.

A necessary condition for the formation of a continuous series of solid solutions of the components K and K′ is that they have the same crystal structure. This requirement greatly limits the possibility of such solutions between elements of B subgroups and, in particular, of subgroup B_1. For elements of subgroup B_2 where the binding is of pronounced covalent character, some well known cases of continuous series of solid solutions are given in Table 31.

Table 31

Continuous Series of Solid Solutions of Elements of Subgroup B_2

Subgroup	Structural type	System
IV B	Coordination	Si – Ge
V B	Layer	As – Sb As – Bi Sb – Bi
VI B	Chain	Se – Te
VII B	Molecular	I – Br*

* The solubility range has been investigated as far as 38 mol % Br.

These cases belong respectively to structures of the coordination, layer and chain types. A wide range of solubility is found in molecular island crystals in the Cl – Br system. A continuous range of solid solutions of elements of subgroup B_1 is formed only in the copper subgroup: Cu – Au, Ag – Au. The corresponding systems are typically metallic.

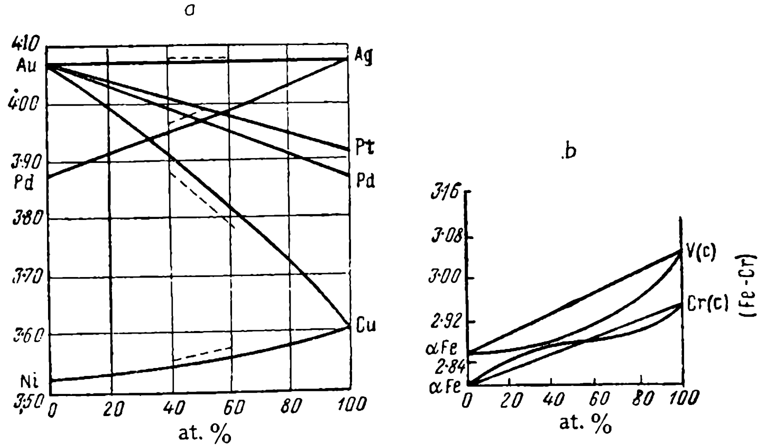


FIG. 87. Concentration dependence of the lattice parameter of the cubic cell of a solid solution with (*a*) face-centred structure, (*b*) body-centred structure.

Change of the Cell Parameters of Solid Solutions with Concentration

In the early days of X-ray studies of solid solutions it was found that the parameters of the unit cell varied linearly with concentration (Vegard's law). Later measurements showed that in most cases the linear dependence was a first approximation. Fig. 87*a* gives experimental data for a variety of metallic solid solutions with face-centred structure and shows that the calculation of the cell parameter from the concentration

$$a = a(c) = a_1 + c\Delta_1 a + c^2 \Delta_2 a +$$

requires the square term also to be taken into account. The second approximation is thus a parabolic dependence. The departures from the linear law (indicated by the dotted line in Fig. 87*a*) may be such that the parabola lies below the straight line (Au–Ag, Pd–Ag, and Ni–Cu) or above it (Au–Cu). In Fig. 87*b* curves are given for several solid solutions of the transition metals which have body-centred structures. The Fe–Cr system has a complex S-shape for the dependence of parameter on concentration.

Local Strains in the Atomic Structure of Solid Solutions

Differences in the atomic sizes of the components leads to local strains in the lattice and these have important effects on the properties of the solid solutions; in particular they lead to changes of the mean lattice parameter. Local strains are shown schematically in Fig. 88 for the two cases where the atoms of the solute are respectively larger and smaller than the atoms of the solvent. The lattice positions

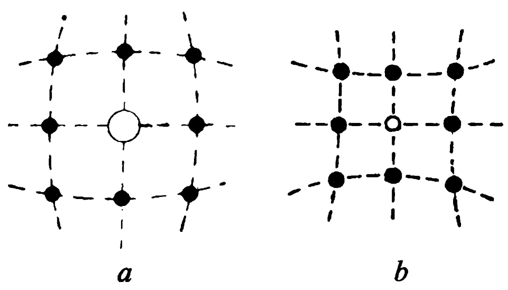


FIG. 88. Local strains in the lattice of a solid solution: solvent atoms (*a*) larger and (*b*) smaller than those of the solute.

occupied by the solute atoms are centres of maximum strain and at them the major part of the strain energy is concentrated.

The amount of lattice strain in a solid solution is determined by the magnitude of the mean square static displacement of the atoms from the equilibrium position u^2_{st}.

From the elastic model of a solid solution it follows that

$$\overline{u^2_{st}} = \gamma c \Delta R^2, \qquad \ldots [6.2]$$

where c is the concentration of solute (assumed small), ΔR the difference in the atomic radii and γ is a numerical factor equal to 7.8 for an f.c.c. lattice and 7.3 for a b.c.c. The magnitude of the static stresses has to be corrected by taking account of the dynamic displacements u^2_d of the atoms from their equilibrium sites as a result of thermal motion in the crystal. For a monatomic cubic crystal

$$\overline{u^2_d} = \frac{3h^2T}{4\pi^2 mk\theta^2}\left[\Phi(x)+\frac{x}{4}\right], \qquad \ldots [6.3]$$

where k and h are the Boltman and Planck constants, m is the mass of the atom, T the absolute temperature, θ the Debye characteristic temperature and $x = \frac{\theta}{T}$ $\Phi(x)$ is the Debye function (see Chapter 14 §1). At sufficiently high temperatures ($x \ll 1$) the expression in square brackets differs little from unity. The characteristic temperature depends on the forces of interaction between particles.

In a solid solution the magnitude of the dynamic displacements can change, either on account of changes in the masses of the atoms or on account of changes in the forces of interaction. The total value of the mean squared displacement of atoms in a solid solution is determined by the sum of [*2*] and [*3*]

$$\overline{u^2} = \overline{u^2_{st}} + \overline{u^2_d}$$

and can be measured by the intensity of X-ray scattering. Investigation of scattering at different temperatures enables us to separate static and dynamic strains. In Fig. 89 experimental data are given on the concentration dependence of static strains in solid solutions of titanium or chromium in nickel. These show that in the alloys there is a linear dependence throughout a fairly wide range of concentration. At the same concentration the amount of lattice strain is greater for alloys of titanium in nickel than for alloys of chromium in nickel because titanium differs more from nickel in its atomic radius than does chromium. Table 32 gives data on the static and dynamic displacements in solid solutions and shows that these displacements are of comparable size.

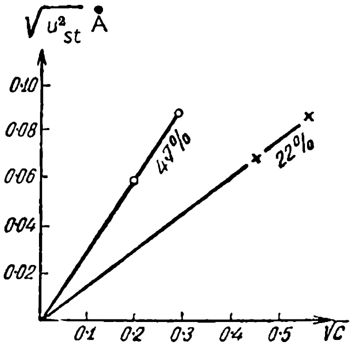


Fig. 89. Concentration dependence of static strains in Ni-based solid solutions containing Ti (0) and Cr (+).

Table 32

Static and Thermal Displacements of Atoms in Some Solid Solutions

Alloy	$\sqrt{\overline{u_{st}^2}}$, Å		$\sqrt{\overline{u_d^2}}$ Å	
		295° K	180° K	90° K
Cu_3Au	0·08	0·14	—	0·09
CoPt	0·07	0·12	—	0·07
NiAu	0·11	0·16	0·13	—
LiMg	0·11	—	0·25	0·16

Ordered Solid Solutions

In solid solutions of simple rational composition A_nB_m there are two possible ways of distributing the atoms: statistical order and disorder. In the former case the probability of finding atoms of different kinds on the different lattice sites is either zero or unity. In the second case the probability is proportional to the concentration. In Table 33 the ordered atomic arrangements (superlattices) which exist in metallic alloys with simple rational composition are given for the three basic types of metallic lattices.

Table 34 gives the symmetry relations between the cell of the superlattice and of the initial lattice, the factor by which its volume changes (this can be obtained from the number of atoms in each unit cell) and the coordination correlations.

In the **b.c.c. structure** we can distinguish two types of structural sites α and β lying at the vertices and centres of the cells. For a disordered alloy with composition AB the probability of finding an A or a B atom on a given lattice site is the same and is one half. These are the structures of the FeAl, CuZn and other alloys at high temperatures. In the fully ordered state there are only A-atoms on the α-sites and B-atoms on the β-sites. The probability of finding an A-atom on an α-site is unity and on a β-site zero. This is the structure of annealed alloys below the transition temperature.

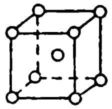
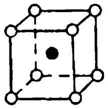
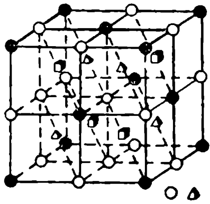
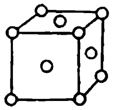
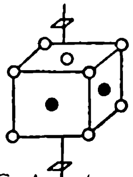
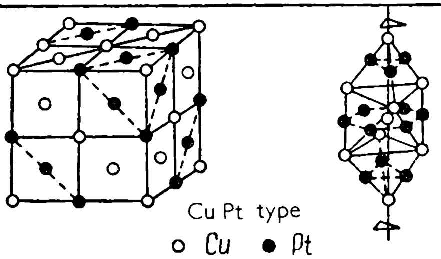
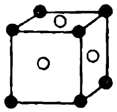
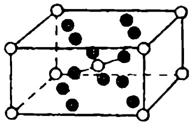
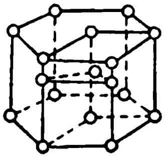
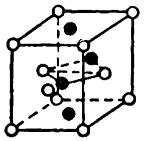
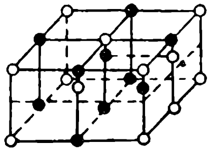
A more complex superlattice is formed in the case of the composition A_3B but it is not difficult to see that by superposing four f.c.c. structures displaced by $\frac{1}{4}$ of the body diagonal, a b.c.c. structure is formed with its cube edge reduced by a factor of two. Conventionally this is written in the form

$$8\Gamma'_c = 4\Gamma''_c = 4(\alpha+\beta+\gamma+\delta) = 4\alpha_3\beta,$$

where α, β, γ, δ, are four families of lattice sites. If three of them are occupied by A-atoms and the fourth by B-atoms we have the superlattice of the Fe_3Al alloy. In this type of structure there is the possibility of forming ternary superlattices with composition A_2BC as indeed exist in the Heusler alloys (Cu_2MnAl, Cu_2MnSn). These alloys are interesting in that, although formed of non-ferromagnetic components they become ferromagnetic when an ordered structure is produced in them by annealing.

There are several types of superlattice based on the **f.c.c. structure**. For the AB composition the simplest ordered structures are formed by a layer-type alternation of A and B atoms such as ABAB . . . along parallel axes of symmetry. By arranging the layers along a fourfold symmetry axis one gets a tetragonal structure of the CuAu

Table 33
Types of Metallic Superlattice

Original structure	Superlattices		
	AB	A_3B	A_4B
 body-centred cubic	 FeAl - CuZn type ○ Fe ● Al	 ○ △ ◪ Fe Fe_3Al type ● Al	
II  face-centred cubic	 Cu Au type ○ Cu ● Au  Cu Pt type ○ Cu ● Pt	 Cu_3Au type ○ Cu ● Au	 Ni_4Mo type ● Ni ○ Mo
III  hexagonal close-packing	 MgCd type ○ Mg ● Cd	 Mg_3Cd type ● Mg ○ Cd	

type, and by arranging them along a threefold axis one gets a rhombohedral structure of the CuPt type. In both these cases ordering is accompanied by a small distortion of the original cubic lattice. For the tetragonal CuAu structure the distortion is measured by the ratio of the axes of the cell $c/a' = 0{\cdot}93$. For the rhombohedral structure we can measure it by the ratio of the body and side diagonals of the rhombohedron; for CuPt this ratio is $c_n/a_n = 1{\cdot}205$ while for a perfect cube it is $\sqrt{3/2} = 1{\cdot}225$. Thus the distortion of the CuPt structure is described by the ratio $\frac{1{\cdot}205}{1{\cdot}225} = 0{\cdot}98$, a value somewhat less than that for CuAu. Alternating the layers along a diad axis does not give a new structure and leads to a superlattice of the CuAu type. The ordered structure with composition A_3B is formed by placing the B-atoms on the vertices and the A-atoms in the centres of the cell faces; in this arrangement the cubic symmetry is not spoiled. This type of ordering is found in a large number of alloys. Composition A_4B has also a possible ordered distribution which is found in Ni_4Mo and Ni_4W; the corresponding superlattice has a tetragonal unit cell with axial ratio $c'/a' = \sqrt{2/5}$.

In the **h.c.p. lattice** there are two types of ordered structure: composition AB has a rhombic superlattice of the MgCd type with a doubled unit cell and composition A_3B has a hexagonal superlattice of the Mg_3Cd type with a unit cell four times larger than that of the original structure.

Comparison of the coordination of disordered and ordered distributions (Table 33) shows that in the majority of cases the ordered form corresponds to an enhancement of the fraction of unlike atoms on nearest neighbour sites. This is so in alloys such as FeAl, Fe_3Al, CuAu, Cu_3Au, Ni_4Mo, MgCd and Mg_3Cd, where the reason for the existence of atomic ordering may be the high strength of the bonds between unlike atoms relative to that of the bonds between like atoms. In CuPt, both in the ordered and disordered states, each atom is surrounded by six neighbours of one kind and six neighbours of the other and the different coordination lies only in the coplanar arrangement of atoms of one type within the ordered structure. Thus the reason for ordering in CuPt is different from that in the alloys listed above.

An interesting type of superlattice—the '*defect lattice*'—occurs in pyrrhotite, which is a sulphide of iron with composition close to FeS. Because there are quadrivalent iron ions, there is usually some deficit of iron as described by the formula $Fe_{1-x}S$ where $0<x<0{\cdot}20$. For compositions with $x<0{\cdot}08$ pyrrhotite is antiferromagnetic, while for $x>0{\cdot}08$ it is ferromagnetic, the maximum ferromagnetism being at $x = 0{\cdot}14$ corresponding to the formula Fe_7S_8. The basic structure of pyrrhotite is hexagonal, of the NiAs type with $a = 3{\cdot}45$, $c = 5{\cdot}67$ Å, $c/a = 1{\cdot}63$ and the unit cell containing 4 atoms (Fig. 90). As was shown by Bertaut, pyrrhotite at room temperature forms a superlattice with a monoclinic cell where $A = 2a\sqrt{3} \approx 11{\cdot}9$ Å, $B = 2a \approx 6{\cdot}865$ Å, $C = 4c \approx 22{\cdot}72$ Å, $\beta = 89° \, 33'$ and the unit cell contains 8 Fe_7S_8 units.

In each cell there are 8 vacant sites (vacancies) arranged in an ordered manner as is shown in Fig. 91. (In this figure only the odd planes of iron are shown; the even planes which do not contain vacancies and also the planes of sulphur atoms are omitted to simplify the drawing.) This superlattice corresponds to the maximum relative separation of vacant lattice sites, the maximum distance between them being $\sim 2a$. The structure described above affords an explanation for the magnetic properties of pyrrhotite: it is antiferromagnetic because there is complete compensation of the magnetic moments of two identical sub-structures formed by the iron atoms in

Table

Types of Superlattice based on

	Type	System	Bravais lattice	Cell parameter	No. of atoms (or sites) in the cell	Space Group
	I. b.c.c.	cubic	Γ'_c	a	2	O_h^9- J_m3m
Super-lattices	FeAl–CuZn	cubic	Γ_c	$a' \approx a$	$1\alpha+1\beta$	O_h^1- P_m3m
	NaTl	cubic	Γ''_c	$a' \approx 2a$	$8\alpha+8\beta$	O_h^7 $Fd3m$
	Fe_3Al	cubic	Γ''_c	$a' \approx 2a$	$12\alpha+4\beta$	O_h^5- $\mathrm{F}m3m$
	II. f.c.c.	cubic	Γ''_c	a	4	same
Super-lattices	CuAu	tetragonal	Γ_t	$a' \approx \frac{a}{\sqrt{2}}+\Delta$ $c' \approx a-\Delta'$	$1\alpha+1\beta$	D_{4h}^1- $\mathrm{P4}/mmm$
	CuPt	rhombo-hedral	Γ_{rh}	$a'_{rh} \approx a\sqrt{2}$ $\alpha_{rh} \approx 60°$	$4\alpha+4\beta$	D_{3d}^5 $R\bar{3}m$

---

† Ferromagnetic alloys of non-ferromagnetic components (Heusler alloys).

34

the Simplest Metallic Lattices

Coordinates of lattice sites	Coordination Arrangement Ordered	Coordination Arrangement Disordered	Examples of superlattices
$(2a)$: 0 0 0; ½ ½ ½	8		
1α: $(1a)$: 0 0 0 1β: $(1b)$: ½ ½ ½	α: 8β β: 8α	$\left(\frac{\alpha}{\beta}\right) 4\alpha+4\beta$	FeAl; CuZn β'-AgZn, β_1-AgCd β_1-AuCd, γ-CuBe
(0 0 0; ½ ½ 0; ½ 0 ½; 0 ½ ½) 8α: $(8a)$: 0 0 0; ¼ ¼ ¼ 8β: $(8b)$: ½ ½ ½; ¾ ¾ ¾	α: $4\beta+4\alpha$ β: $4\alpha+4\beta$	$\left(\frac{\alpha}{\beta}\right) 8\alpha+8\beta$	
(0 0 0: ½ ½ 0: ½ 0 ½: 0 ½ ½)× 4β: $(4a)$: 0 0 0 $4\alpha_1$: $(4b)$: ½ ½ ½ $8\alpha_2$: $(8c)$: ¼ ¼ ¼: ¾ ¾ ¾	α_1: $8\alpha_2$ α_2: $4\alpha_1+4\beta$ β: $8\alpha_2$	$\left(\frac{\alpha}{\beta}\right) 6\alpha+2\beta$	Fe_3Al; β-Fe_3Si β-Cu_3Al Cu_2MnAl; Cu_2MnSn†
$(4a)$: (0 0 0; ½ ½ 0; ½ 0 ½; 0 ½ ½)	12		
1α: $(1a)$: 0 0 0 1β: $(1b)$: ½ ½ ½	α: $8\beta+4\alpha$ β: $8\alpha+4\beta$	$\left(\frac{\alpha}{\beta}\right) 6\alpha+6\beta$	CuAu‡ MnAu
$1\alpha_1$: $(1a)$ 0 0 0 $3\alpha_2$: $(3e)$ 0 ½ ½; ½ 0 ½; ½ ½ 0 $1\beta_1$: $(1b)$: ½ ½ ½ $3\beta_2$: $(3d)$: ½ 0 0; 0 ½ 0; 0 0 ½	α_1: $6\beta_2+6\alpha_2$ α_2: $6\alpha_1+6\beta_1$ β_1: $6\alpha_2+6\beta_2$ β_2: $6\alpha_1+6\beta_1$	$\left(\frac{\alpha}{\beta}\right) 6\alpha+6\beta$	CuPt

---

‡ The CuAu superlattice is usually described as face-centred.

Table

Types of Superlattice based on

	Type	System	Bravais lattice	Cell parameter	No. of atoms (or sites) in the cell	Space Group
Super-lattices	Cu_3Au	cubic	Γ_c	$a' \approx a$	$3\alpha+\beta$	O_h^1– *Pm*3*m*
	Ni_4Mo	tetragonal	Γ_t'	$a' \approx a\sqrt{\frac{5}{2}}$ $c' \approx a$ $\frac{c'}{a'} = \sqrt{\frac{2}{5}} =$ $= 0{\cdot}633$	$8\alpha+2\beta$	C_{4h}^5– *J*4/*m*
	III. h.c.p.	hexagonal	Γ_h	a, c $\gamma = 120°$ $c/a=$ $= \sqrt{\frac{8}{3}} =$ $= 1{\cdot}633$	2	D_{6h}^4 *C*6/*mmc*
Super-lattices	MgCd	rhombic	Γ_0	$a' \approx a$ $b' \approx a\sqrt{3}$ $c' = c$	$2\alpha+2\beta$	D_{2h}^5 *Pmcm*
	Mg_3Cd	hexagonal	Γ_h	$a' \approx 2a$ $c' \approx c$	$6\alpha+2\beta$	D_{6h}^4– *C*6/*mmc*

34 (*cont.*)

the Simplest Metallic Lattices

Coordinates of lattice sites	Coordination		Examples of superlattices
	Arrangement		
	Ordered	Disordered	
1β: (1α): $0\,0\,0$ 3α: $(3c)$: $0\,\frac{1}{2}\,\frac{1}{2}$: $\frac{1}{2}\,0\,\frac{1}{2}$: $\frac{1}{2}\,\frac{1}{2}\,0$	α: $8\alpha+4\beta$ β: 12α	$\left(\frac{\alpha}{\beta}\right) 9\alpha+3\beta$	Cu_3Au; Cu_3Pt; Pd_3Fe; Ni_3Mn; Ni_3Fe; Ni_3Pd Cu_3Pd; Au_3Mn*
$(0\,0\,0;\ \frac{1}{2}\,\frac{1}{2}\,\frac{1}{2})\times$ 2β: $(2a)$: $0\,0\,0$ 8α: $(8h)$: $\pm(xyo;\ \bar{y}xo)$ $x = 0{\cdot}4$; $y = 0{\cdot}2$	α: $9\alpha+3\beta$ β: 12α	$\left(\frac{\alpha}{\beta}\right) 9{\cdot}6\alpha+$ $+2{\cdot}4\beta$	c/a Ni_4Mo 0·623 Ni_4W_0 0·620
$0\,0\,0;\ \frac{1}{3}\,\frac{2}{3}\,\frac{1}{2}$	12		
2α: $(2e)$: $0\,y\,\frac{1}{4}$; $0\,\bar{y}\,\frac{3}{4}$ 2β: $(2f)$: $\frac{1}{2}\,y\,\frac{1}{4}$: $\frac{1}{2}\,y\,\frac{3}{4}$	α: $8\beta+4\alpha$ β: $8\alpha+4\beta$	$\left(\frac{\alpha}{\beta}\right) 6\alpha+6\beta$	MgCd AuCd
$(0\,0\,0$: $\frac{1}{3}\,\frac{2}{3}\,\frac{1}{2})\times$ 2β: $(2c)$: $0\,0\,0$ 6α $(6h)$: $\frac{1}{2}\,0\,0$: $0\,\frac{1}{2}\,0$: $\frac{1}{2}\,\frac{1}{2}\,0$	α: $8\alpha+4\beta$ β: 12α	$\left(\frac{\alpha}{\beta}\right) 9\alpha+3\beta$	Mg_3Cd; $MgCd_3$; Co_3Mo; Co_3W

* These lattices have tetragonal compression.

the ideal composition FeS. As a result of the ordered arrangement of vacancies in a crystal of composition Fe_7S_8, the substructures are no longer identical, and do not give full compensation. This leads to ferromagnetism similar to that in ferrites.

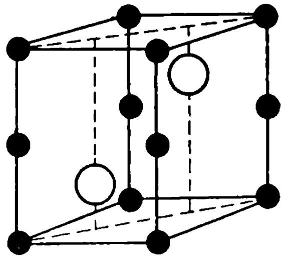


FIG. 90. Basic unit cell.

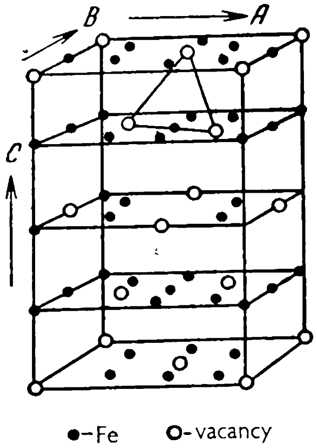


FIG. 91. Superlattice of vacant sites in pyrrhotite Fe_7S_8.

Interstitial Solid Solutions

In the **iron-carbon system** the solubility of carbon in γ-iron may be as high as 8%. The solid solution of carbon in γ-iron is called austenite and, because of the great practical importance of the phase in quenching processes, its structure has been extensively studied. In austenite, as in γ-iron, the iron atoms form an f.c.c. lattice in the octahedral gaps of which the carbon atoms are randomly distributed. This was recently shown by neutron spectroscopy (V. N. Bykov) of austenite alloyed with manganese. Here the neutron scattering by the iron atoms is weakened by the presence of the manganese, the scattering amplitude of which is negative. The dimensions of carbon atoms exceed the dimensions of the octahedral voids, so putting carbon atoms into interstitial positions in the γ-iron lattice is accompanied by an increase in the lattice parameter and density of austenite and also by the development of local static strains in the lattice. If the octahedral voids were all occupied by atoms of the second component the corresponding phase would have the composition AB and be an ordered structure of the NaCl type. Many transition metals form compounds of this type with carbon, nitrogen and oxygen. These compounds include TiC, TiN and TiO and are called *interstitial phases*, but this is, strictly speaking, incorrect since they are metals in the majority of cases and are not f.c.c. However the interstitial phases have interesting physical properties: carbides and nitrides are very hard substances with high melting points and good electrical conductivity; some of them, for example NbC, NbN and ZrN, are superconducting with a high temperature of transition to the semiconducting state.

The explanation of the nature of interstitial phases is very interesting. They arise because of the directed octahedral bonds such as are formed by the mutually-perpendicular p-valence electron clouds of carbon, nitrogen and oxygen. Formation of octahedral bonds by transition metal atoms is also well known on the example of the complex octahedral ions. The atomic radius of carbon in these compounds corresponds, as was shown in Chap. 4 §1, to electron deficiency bonding with multiplicity less than unity, while in oxides the bonds are largely ionic because of the great difference in electronegativity between atoms of oxygen and of the metal.

3. The Structures of Some Compounds

From the very wide subject of the crystal chemistry of compounds we shall consider the structures of only a few groups of compounds which are of greatest interest for the theory of solids.

The Structure of Alloys (Electron Compounds)

In binary alloys of metals of the I B subgroup, Cu, Ag and Au, with elements of the other B-subgroups, the number of valence electrons n_e in the alloy increases with increasing atomic concentration of the second component. This leads to an increase in the concentration of the electron gas, as measured by the ratio n_e/n_a, and to a change in the internal energy of the alloy. As was shown at first empirically by Hume-Rothery, a change in the electron concentration of the alloy is accompanied by the appearance of structurally similar phases which are known by the name of 'electron compounds'. We shall consider the structures of typical such electron compounds taking the Cu – Zn system as an example. In this system there are 5 phases, α, β, γ, ε and η; their structures are shown in Fig. 92. The α-phase is a solid solution of zinc in copper with limiting concentration around 38 at. %. Zn has an f.c.c. structure. The β-phase, which exists at concentrations of about 50 at. % Zn, is b.c.c. The γ-phase is also cubic but with a more complex structure in that it has 52 atoms in its unit cell (Fig. 93). This structure is formed from the β-phase structure as follows: pick out of the β-phase structure a cube with edge parallel to the cell edge and containing 27 unit cells and 54 atoms; from this large cell take away 2 atoms according to the body centring law (from the cube corner and centre); then fill the vacancies thus formed by moving in the surrounding atoms. As a result we arrive at the γ-phase

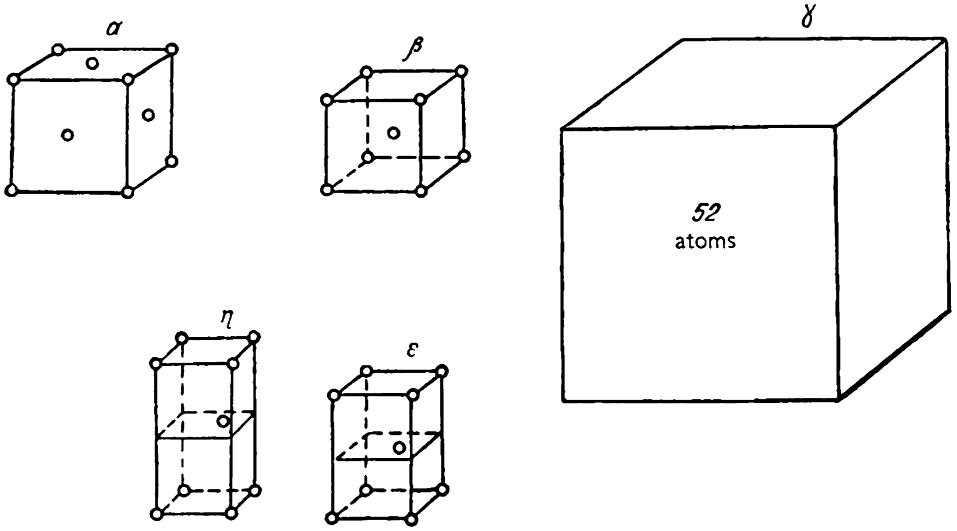


FIG. 92. Cells of electron compounds (α-, β-, γ-, ε-phases).

Table 35

Electron Compounds

n_e/n_a	3/2		21/13	7/4
Phase	β	β'	γ	ε
Symmetry	cubic	cubic	cubic	hexagonal
Type	δ-Mn	β-Mn	giant cell	close packing
Compounds	AgMg AgZn AgCd AuMg AuZn AuCd CuBe CuZn Cu_3Al Cu_3Ga Cu_5Sn	Ag_3Al Au_3Al Cu_5Si	Ag_5Zn_8 Ag_5Cd_8 Ag_5Hg_8 Au_5Zn_8 Au_5Cd_8 Cu_5Zn_8 Cu_5Cd_8 Cu_5Hg_8 Cu_9Al_4 Cu_9Ga_4 Cu_9In_4 $Cu_{31}Si_8$ $Cu_{31}Sn_8$	$AgCd_3$ Ag_3Sn Ag_5Al_3 Ag_5In_3 Cu_3Ge Cu_3Si
	CoAl FeAl NiAl	$CoZn_3$	Co_5Zn_{21}	$FeZn_7$
			$Cu_7Zn_4Al_2$	

structure (Fig. 93) which can be considered as a distorted β-phase superlattice. The ε-phase is h.c.p. with axial ratio $c/a = 1{\cdot}55$. The η-phase is a solid solution of copper in zinc and is a stretched h.c.p. with $c/a = 1{\cdot}85$. The first four phases form a group of electron compounds. Phases with analogous cells and structures are found in a large number of alloys, as can be seen from Table 35 where some of them are listed.

In the majority of cases, electron compounds have well marked solubilities, and the chemical formulae give their approximate compositions. If one element is replaced by a homologous one e.g. Cu – Ag – Au or Mg – Zn – Cd – Hg and so on, the stoichiometric composition of the corresponding compound is unchanged. If the second component is replaced by one of higher valency e.g. Zn – Al – Sn – Sb, this leads to a regular movement of the region of stability of the electron compound on the composition diagram and a marked change in stoichiometric composition. Thus,

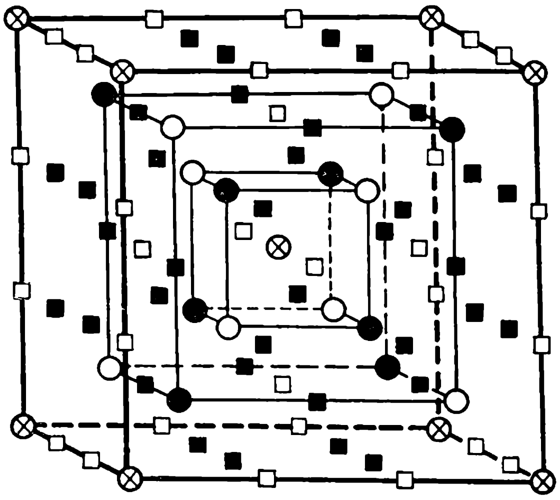


FIG. 93. Structure of the γ-phase.

in the case we were considering, the β-phases are described by the formulae CuZn, Cu_3Al, Cu_5Sn all of which have the same concentration of valence electrons, viz. $n_e/n_a \simeq 3/2$. At concentration $n_e/n_a \simeq 21/13$, compounds are found with the γ-phase structure and also having different stoichiometric compositions (Cu_5Zn_8, Cu_9Al_4, Cu_3Sn_8). Finally the ε-phase appears at electron concentration $n_e/n_a \simeq 7/4$ ($AgCd_3$, Ag_5Al_3, Ag_3Sn).

As a rule the stoichiometric formulae of the electron compounds do not correspond to the number of lattice sites in the unit cell since there is a statistical distribution of atoms of the different kinds. When $n_e/n_a = 3/2$ there is sometimes a second structural type—the β'-phase which has a cubic lattice with 20 atoms in the unit cell. The structures of the β- and β'-phases correspond to those of the two modifications of manganese, δ-Mn and β-Mn. The limit of solubility of the α-phase occurs when $n_e/n_a = 1{\cdot}36$, and the range of solubility falls as the valency of the second component increases (Fig. 94).

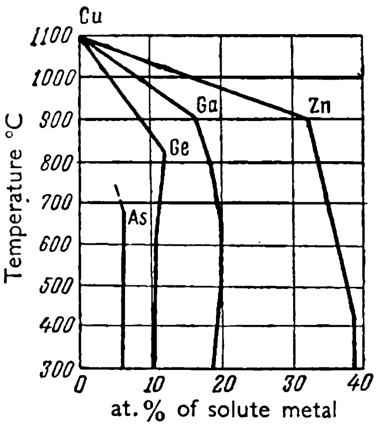


FIG. 94. Dependence of the solubility range of the α-phase on the valency of the second component.

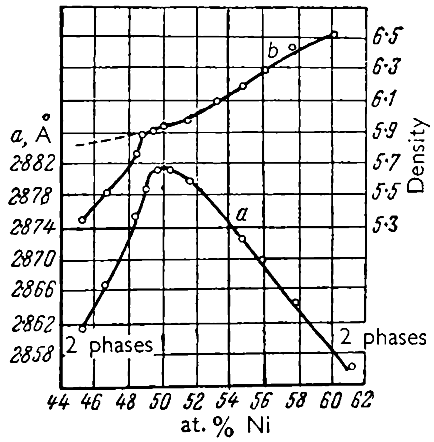


FIG. 95. Dependence of the lattice period and density of the β-phase of Ni - Al on concentration.

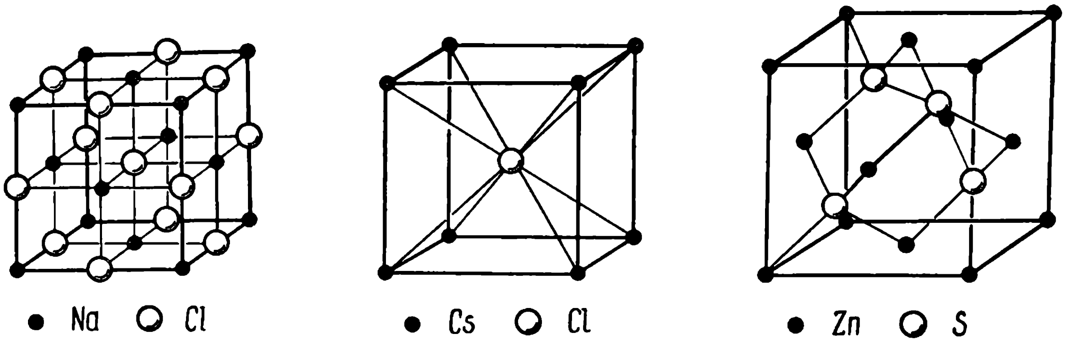
Electron compounds are also formed when transition metals belonging to the I B subgroup (the ferromagnetic triad and Mn, Pd, Rh and Pt) are melted together with elements of the B subgroups. Some compounds of this type are given in Table 35. The law of electron concentrations is observed for these compounds if the number of electrons given to the electron gas by the transition metal atoms is taken as zero. This difference between the transition metals and the metals of the I B subgroup is connected with the fact that the former have unfilled bands.

Interesting results on the structure of the β-phase of the Ni-Al system were obtained by X-rays. This phase exists in the 40-55 at. % Al range; the variation of the lattice parameter in this range is shown in Fig. 95 (curve *a*). At first, as the aluminium increases in concentration and replaces nickel atoms, the lattice parameter increases since the atomic radius of aluminium is greater than that of nickel; the density meantime falls (curve *b*). After the concentration passes 50% further increase in aluminium content leads to reduction of the lattice parameter and to a steeper fall of the density (a similar phenomenon is found also in β-CoAl). Analysis of this strange occurrence in β-NiAl showed that in the Al-rich region an increase in Al content proceeded not by replacement of nickel atoms but by removal of some of the nickel atoms to leave vacant sites in the lattice ('subtraction solution'). The number of aluminium atoms in the cell remains constant and equal to unity but the number of nickel atoms is progressively reduced from unity to 0·83. The total number of atoms in the unit cell of β-NiAl is thus reduced and this phase can be considered amongst those which have a variable number of atoms in their unit cells. In a subtractional solid solution the electron concentration increases more slowly than in a substitutional solid solution; this leads to an increase in the width of the region in which the phase exists.

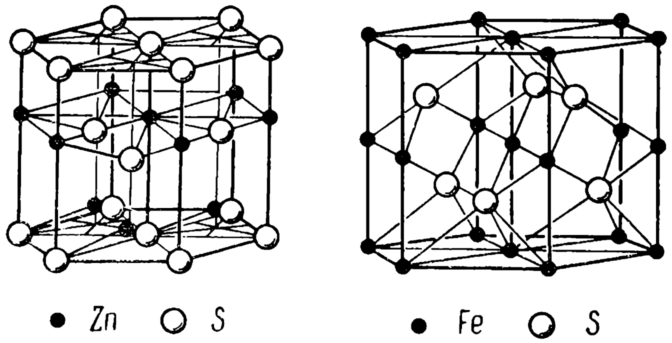
Diamond-like Compounds

Consider a binary compound formed from elements of the B-subgroups which are symmetrically placed on either side of the IV B subgroup, that is from elements $B_1 = B^{IV-q}$ and $B_2 = B^{IV+q}$ where $q = 0, 1, 2$. The corresponding compounds exhibit simple stoichiometry and structure. Stoichiometry is described by the formula $B^{IV-q}\,B^{IV+q}$ which shows that the total number of valence electrons per pair of component atoms is $(4-q)+(4+q) = 8$ thus filling all the *s*-*p* valence levels. When the difference in electronegativity of the components is small (q small) the structure of the compound is determined by tetrahedral sp^3 hybridisation which is typical of structures of elements of the IV B subgroup.

The simplest structures with tetrahedral bonding are those of the two modifications of zinc blende: the cubic sphalerite (α-ZnS) and the hexagonal wurtzite (β-ZnS) (Fig. 96). These structures can be described by the numerical symbols $(\infty, 0)$ and $(1, 1)$. The structure of sphalerite goes over to that of diamond if both components



NaCl CsCl ZnS, sphalerite



ZnS, wurtzite FeS

FIG. 96. Structure of ionic crystals with composition AB.

are of the same kind. There is only one compound known for which $q = 0$; this is SiC (silicon carbide), which forms a variety of modifications, the lattices of which are held together by tetrahedral bonds but have a high degree of stratification. The commonest are the hexagonal modifications: six and four-layer with numerical symbols (3, 3) and 2, 2). One also encounters rhombohedral 15-layer crystals with numerical symbol (2, 3) 2, and more complex ones. In individual cases crystals have been found with a very high degree of stratification into hundreds of atomic layers (superperiodicity)—a phenomenon which finds its explanation in the spiral growth of crystals (Chapter 13). As the difference in electronegativity of the components increases ($q \to 3$) the bonds become fundamentally ionic and structures of ionic type appear. The succession of morphotropic transitions in these compounds can be described as a function of the difference in electronegativities by the following scheme.

Structural type: Graphite — BN
Diamond — α-ZnS — β-ZnS — NaCl
sphalerite wurtzite rock salt

$0 \to q$ (difference in electronegativity of the components of the compounds).

For compounds formed by elements of the second period—BN, BeO (wurtzite structure) and LiF (NaCl structure) we may note the following properties: Boron nitride has a layer structure formed from planes of hexagonal atomic nets based on sp^2 hybridisation and similar to the nets in the graphite structure. Trigonal hybridisation which takes place in boron nitride is typical of a series of the compounds of both boron and nitrogen. Boron and nitrogen can also form tetrahedral bonds based on sp^3 hybridisation, as is shown by the existence of the tetrahedral ionic complexes BF_4^- and NH_4^+ which form, for example, KBF_4 and NH_4F. This leads us to suppose that there should be a second modification of boron nitride with the diamond structure and recently crystals have been made of a cubic modification of boron nitride (borazon) with properties close to those of diamond.

The Structures of the Simplest AB and AB_2 Compounds

In compounds with composition AB (Fig. 96) the coordination number k for the atoms of both components is the same. In a cubic structure of the CsCl type it is 8 and the coordination polyhedron, a cube, is the same for both types of atoms. The structures of a series of inorganic compounds can be considered as close-packed, and formed by the particles with the largest volume (in most cases these are the anions), the gaps between them being filled by the cations according to a rule which is typical of the structural type. Thus in the NaCl structure, the chlorine ions form an f.c.c. structure all the octahedral voids of which are occupied by sodium ions. The coordination polyhedron is an octahedron and is the same for both types of ions.

The hexagonal lattice of iron sulphide (or nickel arsenite) is similar to the NaCl structure. The difference lies in the fact that sulphur atoms form a hexagonal close-packed lattice in the octahedral voids of which the iron atoms are distributed. The coordination number in this structure is also 6, but here the coordination polyhedra are different for the different types of atom: the iron atoms are situated in the centres of the octahedra while the sodium atoms occupy the centres of trigonal prisms (Fig. 90).

The lattices of both polymorphic modifications of zinc sulphide, the cubic sphalerite (α-ZnS) and the hexagonal wurtzite (β-ZnS) are constructed on a similar principle. The sulphur atoms in these structures form cubic and hexagonal close-packing, while the zinc atoms occupy half of the tetrahedral voids. The coordination polyhedra in both structures are the same for both types of atoms and are tetrahedra. In the boron nitride lattice the coordination number is 3; the coordination polyhedra for both atoms is the same and, corresponding to the layer character of the structure, are similarly situated triangles.

Fig. 97 shows the structures of some compounds with composition AB_2. In these compounds the coordination numbers for the two types of atoms are in the ratio 2:1 (if nearest neighbours are atoms of different kinds).

In the cubic structure of fluorite CaF_2 the calcium ions form up in cubic close packing, all tetrahedral voids of which are occupied by fluorine ions. The coordination polyhedron for the calcium atoms is a cube and for the fluorine atoms is a tetrahedron. In the hexagonal lattice of molybdenite MoS_2 (Fig. 270*c*) the molybdenum atoms form an h.c.p. array, the tetrahedral voids of which are occupied by sulphur atoms. The atoms of molybdenum are situated in the centre of trigonal prisms formed by the sulphur atoms. The sulphur atoms are situated inside trigonal prisms formed by molybdenum and sulphur atoms and are displaced on the side of the three nearest molybdenum atoms.

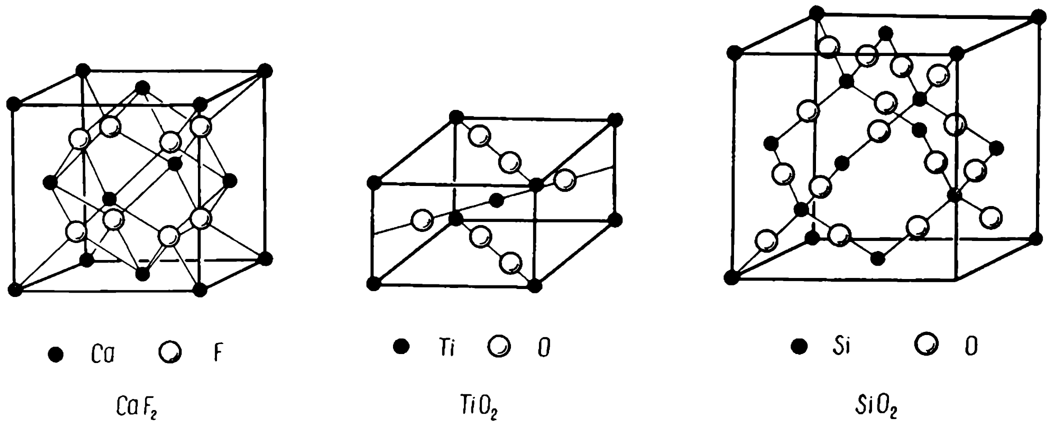


FIG. 97. Structure of ionic crystals with composition AB_2.

In the hexagonal structure of CdI_2 (Fig. 270*b*) the iodine atoms form an h.c.p. array, one half of the octahedral voids of which are occupied by atoms of cadmium. The cadmium atoms are situated in the centres of octahedra and the iodine atoms are inside the octahedra formed by the atoms of cadmium and iodine and are displaced towards the three nearest cadmium atoms. In the tetragonal structure of titanium dioxide (rutile) (Fig. 97) the titanium atoms form a body-centred structure and are surrounded by the oxygen atoms arranged in octahedra. The oxygen atoms are situated in the centres of isosceles triangles formed by the titanium atoms.

In the cubic lattice of pyrites (FeS_2) the iron atoms form a face-centred structure, the octahedral spaces of which are occupied by pairs of sulphur atoms. The sulphur atoms are situated inside tetrahedra, three vertices of which are occupied by iron atoms while the fourth vertex is occupied by a sulphur atom. The iron atoms are octahedrally surrounded by six sulphur atoms. In the cubic lattice of cuprite (Cu_2O) the copper atoms form a body-centred cubic structure, one quarter of the tetrahedral spaces of which are occupied by oxygen atoms. The oxygen atoms form four tetrahedral bonds while the copper atoms form two linear bonds. The coordination numbers of these atoms are respectively 4 and 2. The same coordination ratios are found also in various modifications of silicon dioxide, which are built up on the basis of silicon-oxygen tetrahedra. The silicon atoms form four tetrahedral bonds while the oxygen atoms form two bonds at an angle. The most symmetrical structure is that of the cubic crystobalite in which the silicon atoms, taken alone, form a diamond type structure (Fig. 97).

In conclusion, let us consider the structure of solid carbon dioxide which has a lattice of the molecular type. The linear triatomic CO_2 molecules are placed on the corners and centres of the faces of a cubic cell (a primitive Bravais lattice). The coordination numbers are respectively 2 and 1. It is not difficult to see the topological similarity of the structures of carbon dioxide and pyrites—if we increase the length of the molecules in the carbon dioxide crystal, without changing the orientation and position of their centres and the dimensions of the unit cell, we arrive at a lattice of the pyrites type with pairs of neighbouring oxygen atoms. In spite of the topological similarity the structures of carbon dioxide and pyrites are different from the point of view of crystal chemistry, for the structure of carbon dioxide is of the molecular island type while the structure of pyrites is of the coordination type. The difference in the structural types of these substances is shown up by differences of their properties.

Ionic Compounds

When the difference in electronegativity between the components of a compound is large, the bonds between the atoms remain heteropolar and, in the limiting case, ionic. Valence electrons are rearranged between the atoms which are thus transformed into ions and as a result of ionic interaction the interatomic distance becomes different from the sum of the atomic radii, $d_{AB} \neq r_A + r_B$, and equal to the sum of the ionic radii, $d_{A^+B^-} = (r_{A^+}) + (r_{B^-})$. To determine the ionic radii we have to know only the ionic radius of one of the components or alternatively the ratio of the ionic radii. The latter can be determined from the molar ionic refractive indices which are proportional to the ionic volumes, or it can be determined with the aid of the formula for the radius of the atomic electron orbits, $R = \frac{C_n}{Z - s}$. Here C_n is a constant determined by the principal quantum number of the electrons of the outermost shell of the ion, Z is the nuclear charge and s is a screening constant which can be calculated from X-ray data or from molar refraction. Numerical values of the radii of some anions can also be determined from the dimensions of the unit cells in those cases when it is possible to assume that the anions are in immediate contact: let us consider a group of isomorphic substances with the NaCl structure (Table 36).

Table 36

Interatomic Distances in Isomorphic Compounds with the NaCl Structure

X	Mg–X	Mn–X	Ca–X
O	2·10	2·24	2·38
S	2·60	2·61	2·84
Se	2·73	2·73	2·96

Table

Crystal-chemical Radii (kX): Atomic (Metallic M_{12}

Table 37a: Correction to coordination no. k

k	Radius, %	
	Ionic	Metallic
12	112	100
8	103	98
6	100	96
4	94	88

Table

Correction to the covalent bond

Multiplicity of the bond
1
2
3

I A	II A	III A	IV A	V A	VI A	VII A	VIII A	VIII A
Li 1^{+} 0·68 K 1·33 M 1·55	Be 2^{+} 0·34 K (1·00) M 1·13							
Na 1^{+} 0·98 K 1·54 M 1·89	Mg 2^{+} 0·74 K (1·38) M 1·60							
K 1^{+} 1·33 M 2·36	Ca 2^{+} 1·04 M 1·97	Sc 3^{+} 0·83 M 1·64	Ti 4^{+} 0·64 3^{+} 0·69 2^{+} 0·78 M 1·46	V 5^{+} 0·4 4^{+} 0·61 3^{+} 0·67 2^{+} 0·72 M 1·34	Cr 6^{+} 0·35 3^{+} 0·64 2^{+} 0·83 M 1·27	Mn 7^{+} (0·46) 4^{+} 0·52 3^{+} 0·70 2^{+} 0·91 M 1·30	Fe 3^{+} 0·67 2^{+} 0·80 M 1·26	Co 3^{+} 0·64 2^{+} 0·78 M 1·25
Rb 1^{+} 1·49 M 2·48	Sr 2^{+} 1·20 M 2·15	Y 3^{+} 0·97 M 1·81	Zr 4^{+} 0·82 M 1·60	Nb 5^{+} 0·66 4^{+} 0·67 M 1·45	Mo 6^{+} 0·65 4^{+} 0·68 M 1·39	Tc M 1·36	Ru 4^{+} 0·62 M 1·34	Rh 4^{+} 0·65 3^{+} 0·75 M 1·34
Cs 1^{+} 1·65 M 2·68	Ba 2^{+} 1·38 M 2·21	La 4^{+} 0·90 3^{+} 1·04 M 1·87	Hf 4^{+} 0·82 M 1·59	Ta 5^{+} (0·66) M 1·46	W 6^{+} 0·65 4^{+} 0·68 M 1·40	Re 6^{+} 0·52 M 1·37	Os 4^{+} 0·65 M 1·35	Sr 4^{+} 0·65 M 1·35
Fr	Ra 2^{+} 1·44	Ac 3^{+} 1·11						

Lanthan

Ce	Pr	Nd	Pm	Sm
4^{+} 0·88 3^{+} 1·02 M 1·83	3^{+} 1·00 M 1·82	3^{+} 0·99 M 1·82	3^{+} (0·98) M	3^{+} 0·97 M 1·81

Actin

Th	Pa	U	Np	Pu
4^{+} 0·95 3^{+} 1·08 M 1·80	4^{+} 0·91 3^{+} 1·06 M 1·62	4^{+} 0·89 3^{+} 1·04 M 1·53	4^{+} 0·88 3^{+} 1·02 M 1·50	4^{+} 0·88 3^{+} 1·01 M

37

or Covalent K_1); Ionic I_6 and van-der-Waals V

37b

multiplicity of the
for C, N etc.)

bond length %
100 86 78

							H 1^- 1·36 K 1^+ 0·00	He V_{12} 1·22
			B 3^+ (0·20) K 0·83 M 0·89	C K 0·77	N 5^+ 0·15 3^+ K 0·74 3^- 1·48	O K 0·73 2^- 1·36	F K 0·71 1^- 1·33	Ne V_{12} 1·60
			Al 3^+ 0·57 K 1·26 M 1·43	Si 4^+ 0·39 K 1·17	P 5^+ 0·35 3^+ K 1·10 3^- 1·86	S 6^+ (0·29) K 1·04 2^- 1·82	Cl 7^+ (0·26) K 0·99 1^- 1·81	Ar V_{12} 1·92
Ni 2^+ 0·74 M 1·24	Cu 2^+ 0·80 1^+ 0·98 M 1·28	Zn 2^+ 0·83 K 1·31 M 1·39	Ga 3^+ 0·62 K 1·27 M 1·39	Ge 4^+ 0·44 2^+ 0·65 K 1·22	As 5^+ (0·47) 3^+ 0·69 K 1·21 3^- 1·91	Se 6^+ 0·35 4^+ 0·69 K 1·17 2^- 1·93	Br 7^+ (0·39) K 1·14 1^- 1·96	Kr V_{12} 1·98
Pd 4^+ 0·64 M 1·37	Ag 1^+ 1·13 M 1·44	Cd 2^+ 0·99 K 1·48 M 1·56	In 3^+ 0·92 K 1·44 M 1·66	Sn 4^+ 0·67 2^+ 1·02 K 1·40 M 1·58	Sb 5^+ 0·62 3^+ 0·90 K 1·41 M 1·61 3^- 2·08	Te 6^+ (0·56) 4^+ 0·89 K 1·37 2^- 2·11	I 7^+ (0·50) K 1·33 1^- 2·20	X V_{12} 2·18
Pt 4^+ 0·64 M 1·38	Au 1^+ (1·37) M 1·44	Hg 2^+ 1·12 M 1·60	Tl 3^+ 1·05 1^+ 1·49 M 1·71	Pb 4^+ 0·76 2^+ 1·26 M 1·75	Bi 5^+ (0·74) 3^+ 1·20 K M 1·82 3^- 2·13	Po	At	Rn
VIII A	I B	II B	III B	IVB	V B	VI B	VII B	VIII B

ides

Eu	Gd	Tb	Dy	Ho	Er	Tu	Yb	Lu
3^+ 0·97 M 2·02	3^+ 0·94 M 1·79	3^+ 0·89 M 1·77	3^+ 0·88 M 1·77	3^+ 0·86 M 1·76	3^+ 0·85 M 1·75	3^+ 0·85 M 1·74	3^+ 0·81 M 1·93	3^+ 0·80 M 1·74

ides

Am 4^+ 0·85 3^+ 1·00 M								

The interatomic distance in each case is half the length of the edge of the unit cell. As the atomic number of the anion is increased from oxygen through sulphur to selenium, the dimensions of the unit cell increase, because the radius of the anion increases. If we replace Mg by Mn the cell dimensions of the oxides increase somewhat; this can be explained by the increase in the radius of the Mn ion as compared with the Mg ion. However in the sulphides, the difference of cell dimensions is very small and in the selenides the dimensions are practically the same. What has happened is that, due to the large radius of selenium, the octahedral spaces in the close-packed lattice of selenium ions are large enough to accommodate easily the ions of magnesium and manganese. If we take this to mean that in the selenide lattices the anions are in contact we arrive at the radius of the divalent ion of selenium $Se^{--} = \frac{2{\cdot}73}{\cdot\sqrt{2}} = 1{\cdot}93 kX$.

From the structure of CaSe we find that the radius of Ca^{++} is 1 03kX and from the structures of CaO and CaS we obtain the radius of oxygen $O^{--} = 1{\cdot}35 kX$ and sulphur $S^{--} = 1{\cdot}81 kX$. From anion contacts in the silicates, Bragg obtained the value of 1·35kX in agreement with that given above. Determination of interatomic distances by means of molar refraction gave a smaller value—1·32kX and determination with the aid of the formula for radii of electron orbits gave a larger value—1·40kX. Hence we see that the uncertainty in the absolute magnitudes of ionic radii may be as large as ±3%, but the greatest interest lies in the sum of the ionic radii and not the magnitude of the individual radii. Tables of empirical ionic radii make it possible to obtain the interatomic spacings in ionic crystals with an accuracy of 1-2%. Using the Goldschmidt Table as a basis, the value of the radius of oxygen is 1·32kX; the Pauling Table gives 1·40kX and in the recent table by Bokii and Belov the mean value is 1·36kX.

Table 37 lists data on crystal-chemical radii for the basic types of chemical bond; metallic, covalent and ionic (based on the Bokii-Belov Table) and also for intermolecular and van der Waals interactions. The table also gives magnitudes of metallic and ionic radii, corresponding to coordination numbers 12 and 6, and covalent radii for monatomic bonds. Tables 37*a* and 37*b* give corrections to the coordination numbers for metallic and ionic bonds and for the multiplicity of covalent bonds.

In the structures of binary compounds considered above, the coordination numbers vary within wide limits, and the coordination polyhedra are correspondingly various types of solid; cube, octahedron, tetrahedron and so on. For the same stoichiometry the type of structure may be essentially different and in many cases is determined by the type of coordination polyhedra. The choice of coordination polyhedra for structures with predominantly covalent bonds is determined by the direction of the bonds, and for structures with predominantly ionic bonds by the ratio of the ionic radii of the cations and anions (see Table 20).

For crystals with predominantly ionic binding the limiting ratios of ionic radii make it possible to determine with accuracy the type of coordination. For compounds of the alkali metals with chlorine the corresponding data are given in Table 38.

The majority of these compounds are isomorphic and have the NaCl structure with octahedral configuration. As the radius of the cation is increased the ratio r_C/r_A increases and eventually becomes greater than 0·73, which corresponds to the initial limit of stability of cubic coordination. If a structure with cubic coordination has less binding energy than an octahedral structure there must be a change of structural type (a morphotropic transition) and in fact such a transition takes place in the CsCl crystal. In the case of RbCl the ratio of ionic radii is 0·82, with an accuracy of a few per cent. This value lies above the boundary of the morphotropic transition but nevertheless the crystal retains a structure with octahedral configuration. Under pressure, RbCl and also RbBr and RbI experience polymorphic transitions. As has been shown by X-ray methods, rubidium chloride has the CsCl structure at pressures above 5000 atmospheres. The reason for this polymorphic transition is the dependence of the packing density of ions in the structure on the ratio of the ionic radii and on the coordination number. As Table 38 shows, the packing fraction which is equal to the ratio of the volume of the ions to the total volume of the crystal, falls markedly as r_C/r_A increases. An increase of the coordination number from 6 to 8, as the data for RbCl and CsCl show, is accompanied by much closer packing and reduction of compressibility. Polymorphic transitions of the rubidium halides, which take place under high pressure, are accompanied by a reduction in volume of about 15%. If we took into account the increase of ionic radii of approximately 3% which is connected with an increase of coordination number from 6 to 8 then the total effect of compressing a structure in a polymorphic transition would be still bigger.

Table 38

Ratios of Ionic Radii, Structural Types, Packing Fractions and Compressibilities for the Chlorides of the Alkali Metals

Compound	r_C/r_A	Structural type	Packing fraction %	Compressibility 10^6 cm^2/kg.
LiCl	0·38	NaCl	78·7	3·50
NaCl	0·54	NaCl	65·7	4·18
KCl	0·73	NaCl	56·0	5·65
RbCl	0·82	NaCl	50·2	7·40
CsCl	0·92	CsCl	68·4	5·90

In the CsCl structure the polymorphic transition is produced by lowering the temperature; the low temperature modification of caesium chloride has the NaCl structure. A complete picture of the polymorphic transitions which take place as a result of changes of the basic thermodynamic parameters is given for the compounds we are considering in the following scheme:

$$\begin{array}{lc} & \xrightarrow{\text{Pressure}} \\ & \text{NaCl type} \rightleftarrows \text{CsCl type} \\ \text{Coordination no.} & 6 \qquad\qquad 8 \\ & \xleftarrow[\text{Temperature}]{} \end{array}$$

Among the compounds in which ionic binding is a large factor we must include also the compounds of metals with the most electronegative elements—fluorine and oxygen. Here knowledge of the ratio r_C/r_A also makes it possible to determine typical positions of the boundaries of morphotropic transitions. Table 39 gives the corresponding data for a series of fluorides and oxides of metals with composition ME_2.

All compounds for which the ratio $r_C/r_A < 0·73$ have the rutile structure and compounds for which $r_C/r_A > 0·73$ have the fluorite structure. Zirconium oxide, which lies near to the transition boundary, does not have the coordination which would follow from the ratio r_C/r_A. An exception of this type may be connected with the influence of the mutual polarisation of the ions which leads to a distortion of the spherical shape of the ions. It seems that zirconium oxide has the cubic fluorite structure at high temperatures but that the low temperature monoclinic modification of ZrO_2 has a distorted fluorite structure.

Table 39

Ratios of Ionic Radii and Structural Types for Fluorides and Oxides of Metals with Composition ME_2

Compound	r_C/r_A	Compound	r_C/r_A	Compound	r_C/r_A	Compound	r_C/r_A
Rutile type				Fluorite type			
MgF_2	0·59	MnO_2	0·39	CdF_2	0·77	ZrO_2	0·66
NiF_2	0·59	RuO_2	0·49	CaF_2	0·80	PrO_2	0·76
FeF_2	0·62	MoO_2	0·52	HgF_2	0·84	CeO_2	0·77
ZnF_2	0·62	PbO_2	0·64	PbF_2	0·99	UO_2	0·80
MnF_2	0·68	TeO_2	0·67	BaF_2	1·08	ThO_2	0·84

Analysis of the crystal chemical data shows that the structure of crystals is determined by the number, shape, size and polarisation properties of their constituent atoms, ions or groups of them and the conditions of their formation. This statement, which in a somewhat different form is due to Goldschmid, is the content of the basic law of crystal chemistry.

Problems

1. Given the length of the edge of the unit cube, find the atomic radius of a metal (*a*) with face-centred and (*b*) with body-centred lattice.
2. Derive a formula for calculating the density of a crystal if its composition and structure are known. Carry out the calculations for tungsten, copper and magnesium.

Chapter 7

COHESIVE ENERGY

1. Cohesive Energy of Molecules and Crystals

Cohesive Energy and Internal Energy

The forces of interaction between particles lead to the formation of molecules and crystals and at the same time help to determine their properties. Thus, for example, the larger the binding energy of molecules the higher are their frequencies of vibration and their heats of dissociation. In crystals, an increase in binding energy is usually accompanied by a rise in hardness, strength and melting point. Binding energy increases with increasing numbers of particles in the system. The binding energy, referred to a gram-molecule of the substance is called the cohesive energy.

The cohesive energy is equal to the difference between the potential energies of he system in its initial and final states

$$U = U_1 - U_2. \qquad \ldots [7.1]$$

The final state is that of equilibrium distribution of the particles of the system, at the absolute zero of temperature. The initial state is that in which the particles are so far apart that the interaction energy between them can be neglected. Putting $U_1 = 0$ we obtain from equation [*1*]

$$U = -U_2. \qquad \ldots [7.2]$$

The cohesive energy of diatomic homopolar molecules is equal to the energy of dissociation D_0, extrapolated to 0° K. To get D_0, the internal energy E required to heat the system to temperature T must be subtracted from the heat of dissociation D measured at that temperature. In the case of thermal dissociation of complex molecules, break-up of the molecule takes place through its weakest bond and the energy of dissociation is then equal to the energy of that bond. As a result of dissociation a molecule breaks up into two fragments which have more or less complex structures, so if we are to compare the cohesive energy with experimental data from thermochemical, spectroscopic and other sources it follows that we must define what is the initial state and what are the 'particles' of the system.

In molecular and homopolar crystals the particles are molecules and atoms and for the initial state we take that of a molecular or atomic vapour. The cohesive energy in these cases is equal to the heat of sublimation L_0 extrapolated to 0° K;

$$U = L_0. \qquad \ldots [7.3]$$

In calculations of the cohesive energy of ionic crystals we take for the initial state, a hypothetical state of an ionized gas containing isolated positive and negative ions. For comparison with experiment we have the heats of formation of the corresponding compounds. In thermochemistry, heats of formation Q correspond to the case where the compound is formed from components in standard states at normal temperature and pressure. For example, the components of the thermochemical reaction which forms sodium chloride are metallic sodium and gaseous molecular chlorine

$$\mathrm{Na}_{\mathrm{met}}+\tfrac{1}{2}(\mathrm{Cl}_2)_{\mathrm{gas}} = \mathrm{NaCl}_{\mathrm{cryst}}+Q. \qquad \ldots [7.4]$$

The cohesive energy of ionic crystals U is not the same as Q but can be calculated from Q (§3).

Cohesive Energy of a Diatomic Molecule

We shall now consider a system formed from particles between which central forces act. The potential energy of interaction of the particles in this case depends only on the distance between them $u = u(r)$. We can think of the potential energy as the sum of the attractive and repulsive potentials

$$u(r) = u(r)_{\mathrm{att}}+u(r)_{\mathrm{rep}}. \qquad \ldots [7.5]$$

To simplify the problem further, we assume that both these potentials can be expressed in the form of powers of the separation distance r^{-p} with different values of the exponent p

$$u(r) = -\frac{a}{r^m}+\frac{b}{r^n} \quad (a, b > 0), \qquad \ldots [7.6]$$

where a and b are positive constants. If this function is to have a minimum, the

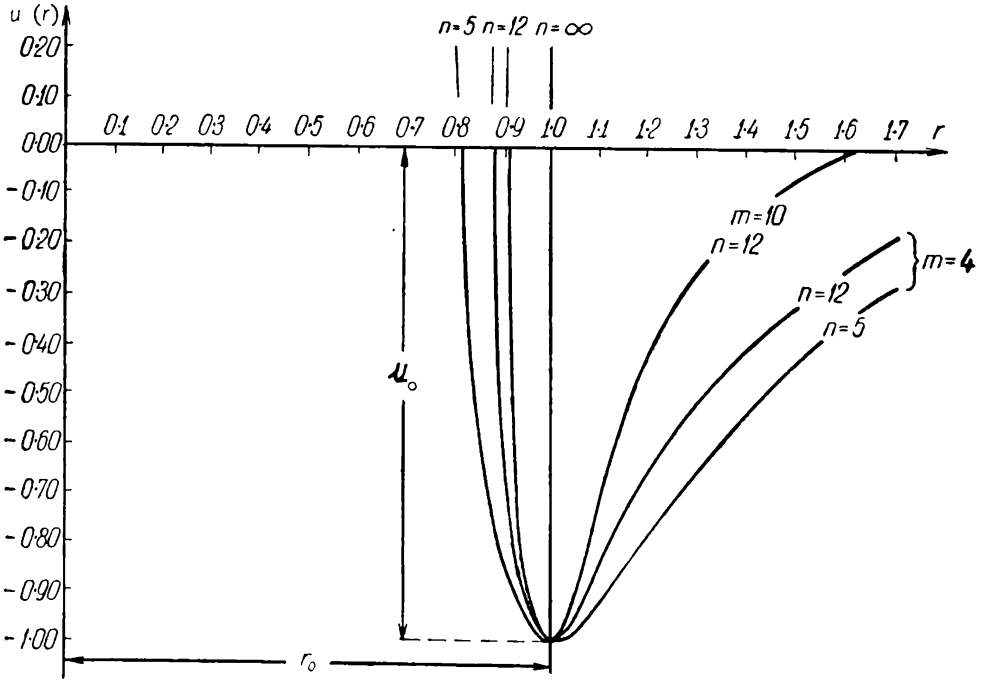


FIG. 98. Potential energy curves for various values of m and n.

exponent of the repulsive potential must be greater than that of the attractive potential: $n>m$. In the case of van der Waals interaction we can take $m = 6$, for ionic interaction $m = 1$. Fig. 98 shows curves of $u(r)$ for various values of m and n. The value $n = \infty$ corresponds to that of absolutely rigid particles and the repulsive potential in this case appears as a vertical straight line parallel to the ordinate axis and at a distance from it given by r_0 equal to the diameter of the particles.

From the condition for an energy minimum $\frac{du}{dr} = 0$ we obtain the equilibrium distance between the particles

$$r_0 = \sqrt[n-m]{\frac{nb}{ma}} \qquad \ldots [7.7]$$

and the cohesive energy of a diatomic molecule

$$u_0 = u(r_0) = -\frac{a}{r_0^m}\left(1-\frac{m}{n}\right). \qquad \ldots [7.8]$$

The first term in [*8*] is the attractive energy and the second is the repulsive energy. If $n \gg m$ the repulsive energy can be considered as a small correction to the attractive energy reducing to zero when $n \to \infty$.

If in [*6*] we replace a and b by parameters u_0 and r_0, which have the obvious physical meanings, we obtain the following expression for the potential energy

$$u(r) = u_0\frac{nm}{n-m}\left[-\frac{1}{m}\left(\frac{r_0}{r}\right)^m + \frac{1}{n}\left(\frac{r_0}{r}\right)^n\right]. \qquad \ldots [7.9]$$

Cohesive Energy of a Crystal

Let us introduce a further simplification and assume that the energy of interaction of two particles does not depend on the presence of other particles. In this case the potential energy of the system is the sum of the energies of interaction of all the pairs of particles and can be expressed in the form of a double summation

$$U = \tfrac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{N} u(r_{ik}). \qquad \ldots [7.10]$$

The summation in [*10*] can be carried out by the method of the 'running vector'. To do this, the vector r_{ik} is first attached to the first particle then to the second and progressively transferred to all the particles. When one end of the vector is attached to a given *i*th particle, the other end runs over all the remaining particles, that is the index k takes all values except $k = i$. Formula [*6*], for a given configuration of particles, then enables us to calculate the energy of the given systems whether they are homogeneous or not.

In homogeneous systems (which include crystals) the configuration and energy states for equivalent particles are the same except for those which lie in a surface layer. Thus, in homogeneous systems the double summation [*10*] can be replaced by a single summation over a 'coordination sphere': consider any particle in the system which does not lie on the surface and take it as centre of a continuously enlarging sphere. Note the moments when this sphere successively cuts the various particles of the system beginning with the nearest, and let $r_1\ r_2 \ldots r_k \ldots$ be the successive radii of co-ordination spheres and $\nu_1\ \nu_2 \dots \nu_k \dots$ the successive coordination numbers for these spheres (ν_k is the number of particles on a sphere of radius r_k so ν_1 is the usual coordination number). It is clear that $r_1 = 2r_a$[†] is the smallest interatomic or intermolecular distance (r_a is the atomic or molecular radius). The cohesive energy of a crystal for which this summation can be carried out is expressed by the formula

$$U = N\sum_{k=1}^{\infty} \nu_k u(r_k). \qquad \dots [7.11]$$

A similar method of summation can be used for crystals which have a high degree of symmetry; for example crystals with structure of the type of simple packing of spheres. We must keep in mine the fact that formula [*1*] includes only the volume energy of the crystal.

2. Cohesive Energy of Molecular Crystals

Structural Sums

The radii of the coordination spheres can be expressed as multiples of the radius of the first sphere r_1; $r_k = r_1\delta_k$ where the coefficient δ_k is uniquely determined by the type of crystal structure and the number of the coordination sphere. The series of coefficients $\delta_1\ \delta_2 \dots \delta_k$ and of coordination numbers $\nu_1, \nu_2 \dots \nu_k$ are characteristic of a given type of structure and do not depend on the size of the unit cell. Substituting in [*11*] the value of the potential energy from [*6*] we obtain the cohesive energy of a crystal in the form

$$U(r_1) = N\left(-\frac{a}{r_1^m}\sum_k \frac{\nu_k}{\delta_k^m} + \frac{b}{r_1^n}\sum \frac{\nu_k}{\delta_k^n}\right) = N\left(-\frac{aA_m}{r_1^m} + \frac{bA_n}{r_1^n}\right), \qquad \dots [7.12]$$

where

$$A_p = \sum_k \frac{\nu_k}{\delta_k^p} \qquad \dots [7.13]$$

are quantities which depend on p and are called structural sums. For a given lattice type the quantities A_p can be calculated for any value of p. Table 40 gives the values of the structural sums for cubic lattices.

Table 40

Structural Sums A_p for Cubic Lattices (Face-centred Γ_c'', Body-centred Γ_c', Simple Cubic Γ_c, and Diamond Γ_d)

p	Jones-Ingham values			Golder's values	p	Jones-Ingham values			Golder's values
	Γ_c''	Γ_c'	Γ_c	Γ_d		Γ_c''	Γ_c'	Γ_c	Γ_d
4	25·338	22·639	16·532	10·183	8	12·802	10·355	6·946	4·331
5	16·968	14·758	10·377	6·289	10	12·311	9·564	6·426	4·092
6	14·454	12·253	8·402	5·113	12	12·132	9·114	6·202	4·039

[†] In the discussion of diatomic molecules above, this distance was denoted by r_0 (see (7)).

For large values of p the value of the structural sum becomes approximately equal to the coordination number. Thus for a b.c.c. lattice the effective value of the coordination number is close to 9. This agrees with calculations of the packing fraction for these lattices (see Table 28, p. 162) according to which the packing fraction of a tetragonal lattice with coordination number $k = 10$ is higher than the packing fraction of a b.c.c. lattice.

Nearest Neighbour Distance and its Dependence on Coordination Number

From the condition for a minimum

$$\frac{dU}{dr_1} = 0 \qquad \ldots [7.14]$$

we obtain the equilibrium distance between particles in the crystal

$$r_1 = r_0 \sqrt[n-m]{\frac{A_n}{A_m}}, \qquad \ldots [7.15]$$

where r_0 is the equilibrium spacing in a diatomic molecule. The spacing between nearest particles in the crystal is different from that in the molecule for, since the quantity A_p becomes smaller as p increases, the distance r_1 is less in the crystal than in the molecule; this can be thought of as the effect of a pressure exerted by the outer coordination spheres on the innermost sphere. Making use of the data of Table 40 we can obtain the dependence of the atomic radius of the particle on coordination number. From Golder's calculations for various values of n and m we obtain the following contractions of the radius as we change from a lattice with coordination number 12 to others with coordination numbers:

k	8	6
$-\Delta r$, %	2·0–2·8	3·5–4·0

The dependence of crystal-chemical radius on coordination number was established empirically by Goldschmidt; the data relevant to metallic and ionic crystals are in Table 37*a*. The contraction of the atomic radius as the coordination number becomes smaller has been measured for metals and these agree with the values calculated for molecular crystals except for diamond-type structures with $k = 4$. In this case, calculation gives a compression which is not much larger than that obtained for $k = 6$ whereas the empirical table gives $\Delta r = 12\%$. However experimental data in this case is scanty. There is a large contraction of the atomic radius (about 9% as the coordination number is changed from 6 to 4) during the polymorphic transition of white metallic tin into grey semiconducting tin. This is due to the change in the type of chemical binding in this transition and the change of the valence state of the tin atom.

Cohesive Energy

Using the condition for a minimum, we can express the energy of a crystal [*12*] in equilibrium in the form

$$U_0 = -\frac{aNA_m}{r_1^m}\left(1-\frac{m}{n}\right) = Nu_0A_m \sqrt[n-m]{\frac{A_m}{A_n}}, \qquad \ldots [7.16]$$

where u_0 is the cohesive energy of a diatomic molecule. Since $A_m > A_n > 2$ the cohesive energy of a crystal is much greater than the cohesive energy of a diatomic molecular gas. The ratio between the energies is

$$q = 2A_m \sqrt[n-m]{\frac{A_m}{A_p}}. \qquad \ldots [7.17]$$

The quantity q can be called the energy strength of the crystal. It is larger in a crystal than in a molecule because the co-ordination number is larger (for a molecule $k=1$). When m and n are large $q \sim 2A_m$ that is, q is about twice the coordination number so for the cubic lattices considered above $q \sim 8$-24. With small values of m and n, q is even larger; thus for a face-centred lattice where $m = 4$ and $n = 5$ we have $q = 75$. This is connected with the broadening of the minimum of the potential curve $u(r)$ and enhancement of the effect of distant coordination spheres. As the magnitudes of the structural sums show (Table 40) a system of particles which interact according to a central force law have a cohesive energy which increases with increase of co-ordination number. Of the four cubic lattices, the structure which is most favoured energetically is that with the closest packing so this calculation supports the basic idea which underlies the theory of close-packing.

This is that, in the absence of directed forces, the thermodynamic condition for equilibrium which is that of minimum free energy (F = min) or, at absolute zero, minimum potential energy

$$U = \text{minimum} \qquad \ldots [7.18]$$

can be replaced by the geometrical condition of minimum volume of the given system.

$$V = \text{minimum} \qquad \ldots [7.19]$$

Energy and Intermolecular Spacing in the Surface Layer of a Crystal

The reduction of energy of interaction of those particles which lie in the surface layer may lead to an increase of the intermolecular spacing. A similar effect was shown by Lennard-Jones (1925). The structural sums and the changes of the molecular radii in a face-centred lattice ($m = 6$, $n = 8$) were calculated by Golder for atoms which lay on the cube face (100) and at various distances below it. He found that in the surface layer there was an increase of the interplanar spacing a by 11%. This effect falls off rapidly and at a distance $5a$ the expansion is only 2%. It follows that electron diffraction measurements of lattice periods in layers a few tens of interplanar spacings thick, will give values 0·8-1·8% above those measured by X-rays in thicker specimens.

Cohesive Energy of Close Packed Spheres

In the various forms of close packing of spheres the packing fraction is the same, so their cohesive energy must be very much the same. The greatest difference of energy of close packing must be expected to be between the cubic and hexagonal types. The structural sums, the lattice compression and the cohesive energy for h.c.p. were calculated by Golder for various values of m and n and various values of the axial ratio c/a. The minimum cohesive energy is that for the structure with axial ratio $c/a = 1{\cdot}633$ which is the nearest to the axial ratio for ideal packing of spheres. With this axial ratio, the cohesive energy for hexagonal close packing is somewhat less in absolute magnitude than that for cubic close packing. Thus in the case of central forces of interaction, the latter structure is energetically more favourable, and this is confirmed by data on the structures of crystals of the majority of the inert elements. X-ray studies show that in the solid state, neon, argon, krypton and xenon have the f.c.c. structure. Only the isotopes of helium have a solid state structure different from f.c.c.

Cohesive Energy and Heat of Sublimation of Molecular Crystals

The total potential of the molecular forces can be expressed in the form

$$u(r) = -\frac{a}{r^6}+be^{-\frac{r}{\rho}}+\varepsilon_0, \qquad \ldots [7.20]$$

where the first term expresses the energy of dipole and dispersion interaction, the second the repulsive energy, and the third the zero point energy. The results of calculations of the cohesive energy of some crystals by means of formula [*20*] are given in Table 41. For comparison, heats of sublimation are also given, these are calculated from the latent heat of evaporation λ ($L = \lambda - RT$) where the second term measures the work of expansion in the transition from the condensed to the gaseous state. Cohesive energies of molecular crystals are very small and are all about 0·5-3 kcal/mole.

Table 41

The Cohesive Energy *U* and Heats of Sublimation *L* for some Molecular Crystals

kcal/mole

Crystal	Ne	Ar	N_2	CH_4
U	0·47	2·09	1·64	2·42
L	0·59	2·03	1·86	2·70

The calculated values of U are in satisfactory agreement with experimental values of L. Any disagreements can be explained by the approximations in the calculations which arise from the difficulty of accurately determining the coefficients in equation [*20*]. The low melting points of molecular crystals are a consequence of the smallness of their cohesive energies.

3. Cohesive Energy of Ionic Crystals

The Madelung Constant

In its simplest form the interaction energy of two ions can be expressed in the form

$$u(r) = \frac{z_1 z_2 e^2}{r}+\frac{b}{r^n}, \qquad \ldots [7.21]$$

where the first term is the electrostatic interaction of the ions considered as point charges, and the second term is the repulsion of the inner, filled electron shells. The coefficients z_1 and z_2 are positive and negative numbers which express the magnitude and sign of the ionic charges. To carry out the summation, we use for the energy of the ionic crystal a formula analogous to [*12*],

$$U(r_1) = N\left(-\frac{z_1 z_2 e^2 A}{r_1} + \frac{B}{r_1^n}\right), \qquad \ldots [7.22]$$

where the quantities A and B correspond to the structural sums [*13*], but with the difference that A in equation [*22*] is expressed by a series of terms with alternating sign coming from the alternation of positive and negative ions in the crystal. The quantity A depends only on the structural type of the crystal and is known as the Madelung constant. Table 42 gives values for the Madelung constant for several simple structural types and also the coordination numbers for the first coordination spheres of ions of each type.

Table 42

The Madelung Constant A and the Coordination Numbers ν_A and ν_B for Simple Types of Structure AB and AB_2

Crystal	CsCl	NaCl Rock Salt	α-ZnS Wurtzite	β-ZnS Sphalerite	CaF_2 Fluorite	α-TiO_2 Rutile	β-TiO_2 Anatase	CdI_2
$\nu_A : \nu_B$	8:8	6:6	4:4	4:4	8:4	6:3		
A	1·763	1·748	1·641	1·638	1·680	1·606	1·600	1·570

For a given stoichiometry the Madelung constant decreases with decreasing coordination number and is not very dependent on the structural type.

As is shown by the examples of the modifications of zinc sulphide and of titanium dioxide, the difference in energy between polymorphic modifications is very small, being only a few tenths of a per cent of the total cohesive energy. This is similar to the energy difference between the cubic and hexagonal forms of close packing and shows the basic difficulty of the theory of polymorphism since the magnitude of the cohesive energy itself can usually be calculated only approximately with an accuracy of several per cent. In such a calculation of the energy of a crystal, certain small terms may be neglected because they do not contribute much to the total energy balance. These small terms may, however, be important in deciding which structural modification is stable at low temperatures.

Cohesive Energy

Using the condition for a minimum of energy,

$$\frac{dU}{dr_1} = \frac{z_1 z_2 e^2}{r_0^2} A - \frac{n}{r_0^{n+1}} B = 0,$$

to determine B,

and eliminating it from [*22*] we obtain the expression for the energy of an ionic crystal

$$U_0 = -\frac{z_1 z_2 e^2 NA}{r_1}\left(1-\frac{1}{n}\right). \qquad \ldots [7.24]$$

If we take it that the ionic charges are known, then in order to calculate the energy from equation [*24*] we need to know *n*—the power of the repulsive potential. This can be obtained if the compressibility of the crystal is known.

The Coefficient of Compressibility

The coefficient of compressibility β is defined as

$$\beta = -\frac{1}{v}\frac{dv}{dp}, \qquad \ldots [7.25]$$

where v is the volume of the body and p is the pressure,

$$p = -\frac{dU}{dv}. \qquad \ldots [7.26]$$

Differentiating [*26*] we obtain

$$\frac{dp}{dv} = -\frac{d^2U}{dv^2}. \qquad \ldots [7.27]$$

From [*25*] and [*27*] we obtain

$$\frac{1}{\beta v} = \frac{d^2U}{dv^2}. \qquad \ldots [7.28]$$

The volume of a gram molecule of the crystal is

$$v = N\gamma r_1^3. \qquad \ldots [7.29]$$

Here N is Avogadro's number, r_1 is the distance between nearest ions and γ is a numerical coefficient of order unity, the magnitude of which is determined by the denseness of packing, that is by the structural type. Since from equation [*29*]

$$\frac{dr_1}{dv} = \frac{1}{3N\gamma r_1^2},$$

we get

$$\frac{dU}{dv} = \frac{dU}{dr_1}\cdot\frac{dr_1}{dv} = \frac{1}{3N\gamma r_1^2}\frac{dU}{dr_1}. \qquad \ldots [7.30]$$

A second differentiation gives

$$\frac{d^2U}{dv^2} = \frac{1}{9N^2\gamma^2 r_1^2}\frac{d}{dr_1}\left(\frac{1}{r_1^2}\frac{dU}{dr_1}\right). \qquad \ldots [7.31]$$

To evaluate the expression in brackets differentiate [*22*]

$$\left(\frac{1}{r_1^2}\frac{dU}{dr_1}\right) = \frac{z_1 z_2 e^2 A}{r^4} - \frac{nB}{r^{n+3}}.$$

Further differentiation gives

$$\frac{d}{dr_1}\left(\frac{1}{r_1^2}\frac{dU}{dr_1}\right) = -\frac{4z_1z_2e^2A}{r^5}+\frac{n(n+3)B}{r^{n+4}} = \frac{z_1z_2e^2(n-1)A}{r^5}, \qquad \ldots [7.32]$$

where in the last operation the value of B has been substituted from [*23*].

Putting [*32*] into [*31*] we obtain

$$\frac{d^2U}{dv^2} = \frac{z_1z_2e^2(n-1)A}{9N^2\gamma^2r_1^2} = \frac{1}{\beta v} = \frac{1}{\beta N\gamma r^3}. \qquad \ldots [7.33]$$

Whence we get, for the coefficient of compressibility

$$\beta = \frac{9\gamma r_1^4}{z_1z_2e^2(n-1)A} = \frac{9vr_1}{z_1z_2e^2(n-1)NA}. \qquad \ldots [7.34]$$

The coefficient of compressibility can be determined by experiment. In this case we can rewrite formula [*34*] to give n

$$n = 1+\frac{9vr_1}{z_1z_2e^2NA\beta}. \qquad \ldots [7.35]$$

The values of n given below were obtained from the compressibility of alkali halide crystals.

Values for n for Alkali Halide Crystals

Crystal	LiF	LiCl	LiBr	NaCl	NaBr
n	5·9	8·0	8·7	9·1	9·5

For the light ions one obtains the smallest values of n, while for the heavier ions $n \sim 8\text{-}10$. As can be seen from equation [*24*], however, inaccuracies in the determination of n have relatively little influence on the energy U since an error of 10% in n gives an error of 1% in U. If the law of repulsion is taken in the exponential form [*3.46*] then in formula [*24*] the correction factor for repulsion is changed to $(1-\rho/r_1)$, and we can determine ρ from the compressibility. The mean value of ρ for the alkali halide crystals is 0·345 Å.

The Born-Haber Cycle. Comparison of the Cohesive Energy with Experimental Data

The cohesive energy U for an ionic crystal, as we have already said, is not equal to the heat of formation Q, but can be obtained from it. To obtain it, we make use of a cyclic process, due to Born and Haber as follows:

$$\begin{tikzcd}
 & \mathrm{Na}^+_{\mathrm{gas}} \arrow[d, "+I_{\mathrm{Na}}"] & + & \mathrm{Cl}^-_{\mathrm{gas}} \arrow[d, "-E_{\mathrm{Cl}}"] & \\
\mathrm{NaCl}_{\mathrm{cryst}} \arrow[ur, "-U_{\mathrm{NaCl}}"] \arrow[dr, "+Q_{\mathrm{NaCl}}"'] & \mathrm{Na}_{\mathrm{gas}} \arrow[d, "+L_{\mathrm{Na}}"] & & \mathrm{Cl}_{\mathrm{gas}} \arrow[d, "+\frac{1}{2}D_{\mathrm{Cl}_2}"] & \ldots [7.36] \\
 & \mathrm{Na}_{\mathrm{cryst}} & + & \frac{1}{2}\mathrm{Cl}_{2_{\mathrm{gas}}} &
\end{tikzcd}$$

In the first part of the process, the crystal is broken up into a gas of ions; this requires the expenditure of an energy equal to the energy of the crystal $-U_{NaCl}$. Next, the ionic gas is transformed into a mixture of the atomic gases, sodium and chlorine; this requires the absorption of energy equal to the ionisation potential of the sodium atom ($+I_{Na}$), and the emission of energy equal to the energy of affinity of the chlorine atom to its electron $-E_{Cl}$. Then the sodium vapour condenses, forming metallic sodium, and the chlorine atoms unite to form diatomic molecules of chlorine. Both processes are accompanied by the emission of energy, equal to the heat of sublimation of the sodium crystal ($+L_{Na}$) and half of the energy of dissociation of molecular chlorine ($\frac{1}{2}D_{Cl_2}$) respectively. The final part of the process consists of the formation of the sodium chloride crystal from its components, considered as being in standard states. In this process energy is emitted equal to the heat of formation $+Q_{NaCl}$. The total energy change in the cyclic process is zero, hence

$$U_{NaCl} = Q_{NaCl} - E_{Cl} + \tfrac{1}{2}D_{Cl_2} + I_{Na} + L_{Na} \qquad \ldots [7.37]$$

All quantities on the right hand side of formula [*37*] are known. Thus we can determine experimentally the energy of an ionic crystal. Table 43 gives data for calculating the energy of various crystals.

Table 43

Cohesive Energy of Alkali Halide Crystals

Crystal	U, erg/mole $\times 10^{-12}$					U, kcal/mole	
	$-Ae^2/r$	c/r^6	$\text{Be}^{-r/\rho}$	ε_0	U	Calculated	Experimental
NaCl	−14·18	−0·20	+1·63	+0·12	−12·63	−183·1	−181·3
KI	−11·30	−0·26	+1·10	+0·07	−10·39	−150·8	−153·8
RbBr	−11·60	−0·28	+1·22	+0·06	−11·60	−153·5	−151·3
CsI	−10·19	−0·46	+1·01	+0·05	− 9·59	−139·1	−141·5

This calculation takes account of the Coulomb energy of interaction of ions which is the basic contribution to the total energy, and also of the van der Waals energy of attraction c/r^6, the repulsive energy, and the zero point energy ε_0. In the final two columns the calculated and experimental values of the crystal energy are given side by side. The largest discrepancy is about 1-2% and indicates good agreement of theory of ionic crystals with experiment.

Fersman, Kapustinsky and others have suggested more accurate formulae, which make it possible to obtain the energy of ionic crystals from the characteristic parameters of ions. One of Kapustinsky's formulae which is obtained from formula [*24*] has the form

$$U_0 = -287{\cdot}2\,\frac{\Sigma\, m z_1 z_2}{r_A + r_C}\left(1 - \frac{0{\cdot}345}{r_A + r_C}\right), \qquad \ldots [7.38]$$

where r_A and r_C are the ionic radii, z_1 and z_2 the ionic charges, m the number of corresponding ions; the summation takes place over all the ions in the chemical formulae.

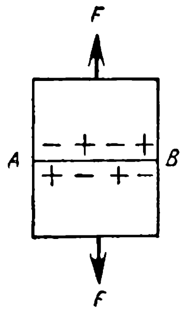


FIG. 99. Calculation of the strength.

Theoretical Strength of a Crystal

The ionic model makes it possible to calculate the theoretical strength of a crystal, that is the resistance to fracture under the influence of external mechanical forces. To obtain the order of magnitude of the strength, consider uniaxial extension of a crystal by the external force F (Fig. 99). The least resistance to fracture is offered by a plane with a chess-board alternation of ions. The line AB is the trace of such a plane. In the first approximation, consider the interaction of nearest opposite charged ions which lie on both sides of a possible cleavage plane. The force pulling these ions apart is $f = \frac{e^2}{r^2}$, and the number of ions in unit area of the plane is $N' = \frac{1}{r^2}$. Hence the resistance of the crystal to fracture is

$$p = \frac{e^2}{r^4} = \frac{23 \times 10^{-20}}{81 \times 10^{-32}} \text{ dynes/cm}^2 = 3 \times 10^{11} \text{ dynes/cm}^2 = 300 \text{ kg/mm}^2 \quad \ldots [7.39]$$

The value thus found is the theoretical strength and gives the upper limit of the strength of a crystal. This limit is not usually attained in practice for a variety of reasons connected with defects in the structure of crystals, and non-uniformity in the way in which the stress is applied. The measured technical strength is always less than the theoretical.

Ionic Model of a Metal

Consider an approximate ionic model for a metal. The crystal structure of a metal is formed of a framework of positively charged atoms the spaces between which are filled with negatively charged electron gas. Electrostatic interaction between the framework and the electrons gives a force of attraction. Assuming that the conduction electrons are uniformly distributed among the ions, we can determine the cohesive energy of a metallic crystal approximately by means of the formula for the energy of an ionic crystal. If z is the valency of the metal, then without taking account of the repulsive forces we have

$$U_{\text{cryst}} = -\frac{2Az^2e^2N}{r_0} \quad \ldots [7.40]$$

where r_0 is the atomic diameter and A is the Madelung constant.

The attractive forces are balanced by the repulsive forces. Now the atomic radii of metals are much greater than the ionic radii of the cations. This can be explained if we assume that in metals the repulsive forces are due, fundamentally, to the motion of the electrons in the electron gas. These forces can be thought of as the electron gas pressure which can be calculated using the formula of kinetic theory

$$pv = \tfrac{2}{3} N\varepsilon = \tfrac{2}{3} K,$$

where ε is the mean kinetic energy of the conduction electrons in the metal and K is the kinetic energy of a gram molecule of electron gas. The kinetic energy of the conduction electrons is much greater than the mean thermal energy ($\varepsilon \gg kT$) and is determined by the density of the electron gas n.

For a monovalent metal $n = \frac{\gamma^{3/2}}{r_0^3}$ where γ is a quantity of order unity which depends on the structural type of the metal. Thus

$$\varepsilon_{\max} \approx \frac{h^2}{8m} n^{2/3} = \frac{h^2}{8mr_0^2}. \qquad \ldots [7.41]$$

If the metal has valency z this formula requires also a factor $z^{2/3}$. The kinetic energy of the conduction electrons is (per gram molecule of metal)

$$K = \frac{h^2 z^{5/3} N}{8mr_0^2}. \qquad \ldots [7.42]$$

In the approximation we are considering the total energy of a metal consists of two terms

$$U = \left(-\frac{a}{r_0} + \frac{b}{r_0^2}\right) N, \qquad \ldots [7.43]$$

where

$$a = 2Az^2e^2; \quad b = \frac{\gamma h^2 z^{5/3}}{8m}. \qquad \ldots [7.44]$$

Atomic Radii of Metals

From the condition for equilibrium of a crystal we obtain

$$\frac{dU}{dr} = \frac{a}{r_0^2} - \frac{2b}{r_0^3} = \frac{1}{r_0^3}(ar_0 - 2b) = 0.$$

Hence we obtain the value of the atomic radius in metallic crystals:

$$r_0 = \frac{2b}{a} = \frac{h^2}{8Ame^2} \cdot \frac{1}{z^{1/3}} = 2{\cdot}62 \frac{\gamma}{A} \frac{1}{z^{1/3}}. \qquad \ldots [7.45]$$

If we take $z = 1$; $A = 1{\cdot}74$ and $\gamma = 1{\cdot}3$ we obtain $r_0 = 2$ Å, which is of the correct order for interatomic spacings in metal crystals.

Formula [*45*] gives the dependence of the atomic radius of a metal on valency. Fig. 100 shows a curve $r_0 = r_0(z)$, calculated by means of equation [*45*]—the points on it are measured values of the atomic radii of metals at the beginning of the fourth period (the theoretical curve has been brought into coincidence with experiment at the value of the atomic radius of potassium).

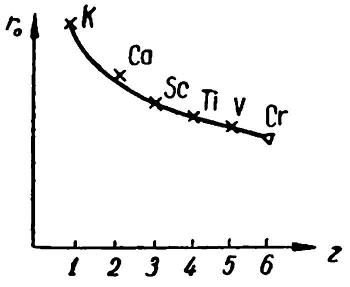


FIG. 100. The dependence of the atomic radius of a metal on the number of valence electrons.

The experimental points lie on the curve and its shape is that of the theoretical curve. This gives support to the conclusion we reached in an earlier chapter viz., that in the transition metals belonging to the subgroup from scandium to chromium the *d*-electrons take part in binding along with the *s*- and *p*-electrons.

Energy of the Alkali Metals

A consequence of the virial theorem is that in equilibrium the kinetic energy of the electrons is equal to half their potential energy. From formula [*43*], the repulsive energy would be a factor two smaller than the attractive energy. Thus the energy of a metallic crystal is equal to the energy to decompose it into ions and electrons

$$U_0 = -\frac{aN}{2r_0} = -\frac{Ae^2N}{r_0} = -\frac{550}{r_0}\text{ kcal/mole} \qquad \ldots [7.46]$$

(here r_0 must be expressed in Angstroms). Table 44 gives values of the energy of the alkali metals calculated by means of equation [*46*].

Table 44

The Energy of Alkali Metal Crystals in kcal/mole

Metal	r_0, Å	E_0	I	L_{calc}	L_{meas}
Li	3·03	182	124	58	39
Na	3·71	142	118	30	26
K	4·62	119	100	19	20
Rb	4·87	113	96	17	18·9
Cs	5·27	104	89	15	18·8

An experimental value of the energy to sublime a metal into its neutral atoms differs from the energy calculated by means of equation [*46*] by the magnitude of the energy of ionisation $L = U_0 - I$. The last two columns are values calculated by means of the semi-empirical formula [*46*] and measured values of the energy of sublimation. Our formula predicts the order of magnitude and the general trend of the changes of energy of sublimation with increasing atomic number of the alkali metal. However the ionic model of a metal gives only rough quantitative agreement with experiment.

The Modulus of Compressibility

To calculate the modulus of compressibility of a metal we may make use of formula [*34*], which was obtained for an ionic crystal, if in it we put $n = 2$ and take twice the value of the Madelung constant. Hence we obtain

$$\beta = \frac{qv_0r_0}{2ANz^2e^2}. \qquad \ldots [7.47]$$

Calculations of the modulus of compressibility for alkali metals by means of this formula give satisfactory agreement with experiment. Thus, for potassium $\beta_{\text{calc}} = 4{\cdot}1 \times 10^{-5}$ cm^2/kg while $\beta_{\text{meas}} = 3{\cdot}6 \times 10^{-5}$ cm^2/kg.

Problems

1. In the formula $v = N\gamma r^3$ which connects the volume of the coordination crystal with r_1 the smallest interatomic distance, there occurs the coefficient γ. Find the

value of this coefficient for the following structures: (*a*) f.c.c. (*b*) b.c.c. (*c*) diamond type (*d*) NaCl type.

2. Find the value of the Madelung constant for a CsCl type crystal (the summation is best carried out over electrically neutral volumes).
3. Calculate the energy for dissociation into neutral atoms of the following ionic molecules (*a*) Na^+Cl^- (*b*) Na^-Cl^+.

Chapter 8

THE ELECTRON THEORY OF CRYSTALS

1. Classical Theory of the Electron Gas

The electron theory of crystals was first worked out for metals. As a first approximation, the conduction electrons can be considered as completely free and one can try to apply the ideal gas laws to them (Drude, 1905). If the mean kinetic energy ε_m of the electrons is equated to the mean kinetic energy of the atoms in the lattice:

$$\varepsilon_m = \tfrac{3}{2}kT, \qquad \ldots [8.1]$$

then the mean velocity of translational motion of the electrons at room temperature ($\varepsilon_m = 0{\cdot}03$ eV) would be $v_m = 10^7$ cm/sec. Hence using the Clapeyron equation we can calculate the pressure of the electron gas p_e: taking, for monovalent metals, the number of free electrons equal to the number of atoms we have

$$p_e = \frac{22400}{V} \text{ atm.,}$$

where V is the specific volume of the metal. For copper the volume of the gm. atom $V = 7{\cdot}1$ cm^3, consequently $p_e \approx 3000$ atm.

To keep the free electrons in the metal we must postulate the existence of a potential barrier on the metal-vacuum boundary, the height of the barrier being $W > \varepsilon_m$. This potential barrier is created by the circumstance that an electrical double layer is formed on the surface of a metal, and that its external face is negatively charged. The width of this double layer is very small, of the order of the atomic radius, or about 10^{-8} cm. Thus energetically, the model of the metal can be considered as a right-angled potential well or box of depth W filled to a depth of the order of ε_m by a layer of free electrons.

The thermal conductivity of the crystal comes from the thermal conductivity of the electron gas and is

$$\kappa = \tfrac{1}{3}\, l v_m n c_1 \qquad \ldots [8.2]$$

where n is the number of free electrons in one cm^3, l is the mean free path and c_1 is the specific heat per electron. On the classical theory $c_1 = 3k/2$ hence

$$\kappa = \tfrac{1}{2}\, l v_m n k. \qquad \ldots [8.3]$$

If the metal is brought into an electric field, the free electrons, still keeping their random thermal motion, begin to take part in directed motion along the lines of force of the field. In the intervals from one collision to the next the electrons travel with an acceleration given by $\frac{eE}{m}$, where E is the strength of the electric field in the metal, e and m are the charge and mass of the electron. If τ is the time of free flight, then the mean additional velocity of the electrons in the direction of the electric field is

$$u = \tfrac{1}{2}\frac{eE}{m}\tau.$$

As a result of this directed motion of the electrons in the metal, a current arises, the density of which is

$$i = neu = \frac{ne^2\tau}{2m}E = \sigma E \qquad \ldots [8.4]$$

and is proportional to the strength of the electric field. Thus this model of the electron gas leads to the basic law of electrical conductivity of metals—Ohm's Law. Assuming that the mean velocity of the electrons v_{m} is much greater than the additional velocity $v_{\mathrm{m}} \gg u$, we obtain $\tau = \frac{l}{v_{\mathrm{m}}}$ and for the electrical conductivity

$$\sigma = \frac{ne^2 l}{2mv_{\mathrm{m}}}. \qquad \ldots [8.5]$$

From [3] and [5] the ratio of the thermal to the electrical conductivity of a metal is

$$\frac{\kappa}{\sigma} = 3\left(\frac{k}{e}\right)^2 T \qquad \ldots [8.6]$$

which is proportional to the absolute temperature if we take for the kinetic energy of the electrons the classical value $mv_{\mathrm{m}}^2 = 3kT$. This expresses another experimental law: the Wiedemann-Franz Law. The numerical value of the coefficient in formula [6] is in good agreement with experiment. Thus for silver at room temperature we have

$$\left(\frac{\kappa}{\sigma}\right)_{\mathrm{meas}} = 1{\cdot}64\times 10^{10} \text{ and } \left(\frac{\kappa}{\sigma}\right)_{\mathrm{calc}} = 1{\cdot}59\times 10^{10}.$$

However a further refinement of the classical theory of the electron gas carried out by Lorentz (he introduced the Maxwellian velocity distribution instead of the mean velocity) did not alter the essentials of the formulae but led, not to an improvement, but to a deterioration of the numerical agreement.

Since with the aid of this model of the electron gas we obtain two basic laws of the metallic state, we may suppose that the model is on the right lines. In fact, as a result of later refinements of the theory on the basis of quantum methods, the general form of the expressions, in particular of [5], are not changed although the content and interpretation of the individual parameters which come into them are essentially different. The necessity for a radical revision of the basis of classical theory became clear when attempts to broaden the region of applicability of the theory and further refine it led to insurmountable disagreements with experiment. We shall show two of these disagreements:

If in formula [5] we replace v_m by an expression involving the temperature, assuming that the kinetic energy of the free electrons is equal to the mean thermal energy, then for the resistivity we get

$$\rho = \frac{2mv}{ne^2 l} = \frac{2\sqrt{mkT}}{ne^2 l} = a\sqrt{T}. \qquad \ldots [8.7]$$

Consequently the resistivity should vary with temperature as $\sqrt{T}$, while experiment gives a variation proportional to T.

Fundamental disagreement is also obtained with the specific heats of metals. If we assume that the number of free electrons is equal to the valency of the atoms of the metal, then for a monovalent metal the number of free electrons is equal to the number of atoms. Thus the electron gas should contribute to the total specific heat its own share which is half the specific heat of the lattice. This leads to an increase of the specific heat by a factor of $1\frac{1}{2}$, that is to a value of 9 calories/gm. atom °C for the atomic heat of a metal, whereas the empirical law of Dulong and Petit gives 6 cal/gm. atom °C. Thus we must accept the fact that the number of free electrons is much less and does not exceed 1-2% of the total number of atoms. On this supposition, the electrical resistivity of the metal calculated from formula [7] comes out much larger than is found by experiment. A large number of free electrons is also necessary to explain the high thermal conductivity of metals. A way out of these contradictions was discovered in the quantum electron theory of crystals.

2. Bringing the Electron Model of a Metal into Agreement with Experimental Data

Consider a model of a metal with the form of a potential box. Take the internal potential of the crystal to be W, the work to remove an electron from a metal to be ϕ,

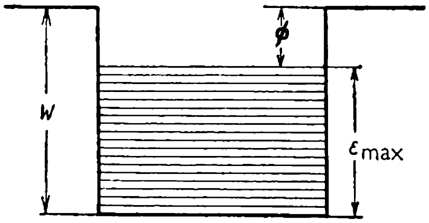


FIG. 101. Potential box energy model of a crystal.

and the maximum kinetic energy of the free electrons to be ε_{max}. The crystal potential is the depth of the potential box (Fig. 101).

The work function or the work to remove an electron from the metal is the distance from the upper surface of the full energy band to the outer boundary of the box. The maximum kinetic energy ε_{max} of the electrons is determined by the height of the upper surface of the layer of electrons above the bottom of the box.

The Internal Potential of a Crystal

If there is an electrical double layer at the metal-vacuum interface, then on taking electrons through this layer from the vacuum into the crystal their kinetic energy increases and on taking them in the opposite direction it is reduced. The change of energy of the electrons in both cases is equal to the internal potential of the crystal W, and this can be measured by diffraction of electron waves. The de Broglie formula in vacuum has the form

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2meV}} = \frac{12{\cdot}22}{\sqrt{V}}\,\text{Å}, \qquad \ldots [8.8]$$

where λ is the electron wave length (in Å) and the accelerating potential difference V is in volts. The velocity of the electron is somewhat increased as it comes into the crystal and its wavelength becomes shorter:

$$\lambda' = \frac{12{\cdot}22}{\sqrt{V+W}}. \qquad \ldots [8.9]$$

The refractive index for electron waves passing from vacuum into a crystal will be

$$\mu = \frac{\lambda}{\lambda'} = \sqrt{1+\frac{W}{V}}. \qquad \ldots [8.10]$$

Refraction of electron waves is observed in electron diffraction experiments, and from them the internal potential for various substances is found to be between 5 and 20 volts.

The Electron Work Function

The work function for electrons—their heat of evaporation—determines the density of thermionic emission current

$$i = aT^2 e^{-\frac{\phi}{kT}}. \qquad \ldots [8.11]$$

Thus by measuring the temperature dependence of thermionic emission we can find the work function for a given metal. Another experimental method for measuring ϕ is based on determining the red cut-off of the photoeffect. The general equation for the photoeffect is

$$h\nu = \tfrac{1}{2}mv^2 + \phi + \phi_1, \qquad \ldots [8.12]$$

where $h\nu$ is the energy of the light quantum absorbed by the crystal, ϕ_1 is the ionisation energy, ϕ is the work function and $\frac{1}{2}mv^2$ is the kinetic energy of a photoelectron. In dielectrics there are no free electrons so the work function $\phi = 0$. Only the ionisation energy ϕ_1 is needed to free a photoelectron. In metals, the ionisation energy for free electrons $\phi_1 = 0$, but the work function $\phi \neq 0$. The equation for the photoeffect in metals is thus

$$h\nu = \tfrac{1}{2}mv^2 + \phi \qquad \ldots [8.13]$$

Table 45

The Work Function ϕ (eV) for Several Metals

Metal	Mo	Rh	Pd	Pt
ϕ from thermionic emission	4·14–4·17	4·58	4·99	6·27
ϕ from the photoeffect	4·15	4·57	4·96	6·30

and from this we can find the work function by measuring the frequency ν_k of the red cut-off. The actual value of ν_k is determined by using the linear relation between the maximum kinetic energy of the photoelectrons and the frequency of the light absorbed (Fig. 102). Table 45 gives values of the work function for several metals determined from thermionic emission and from the photoeffect.

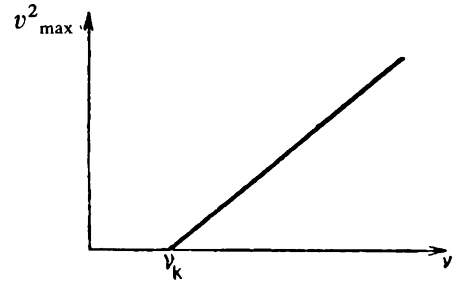


FIG. 102. The red cut-off of the photoeffect.

The Maximum Kinetic Energy of the Free Electrons

Comparison of the magnitudes of the lattice potential and of the work function shows that the filled energy layer is filled to quite a considerable depth, and that its upper limit $\varepsilon_{\max}$ corresponds to an electron energy of from 2-12 eV. Table 46 gives values of the internal potential and the maximum energy of free electrons for several metals. From these it is evident that the potential well in a metal is filled to half or more of its total depth.

Table 46

The Internal Potential W and the Maximum Energy $\varepsilon_{\max}$ of Conduction Electrons for Several Metals (eV)

Metal	Li	Na	K	Cu	Ag	Au	Ca	Al
W	6·9	5·0	3·9	11·1	10·2	10·3	7·5	14·7
$\varepsilon_{\max}$	4·7	3·1	2·1	7·0	5·5	5·5	4·3	11·7

The results show that the kinetic energy of the conduction electrons in a metal is about a hundred times greater than their mean thermal energy. According to the classical theory of the electron gas, the potential well should be filled to only something of the order of $\frac{1}{1000}$ - $\frac{1}{100}$ of its total depth. The classical theory cannot explain the large kinetic energies of electrons corresponding to temperatures of the order of tens of thousands of degrees.

Experiments on the scattering of X-rays by conduction electrons in beryllium show an extraordinary large broadening of the Compton lines. From the breadth of these lines the velocities of motion of the conduction electrons can be calculated and turn out to be about ten times greater than the velocities which the electrons would have if they had a Maxwellian distribution of energy and were in thermal equilibrium with the lattice.

These results are supported by investigations on the annihilation of positrons in metals: fast positrons, which enter the crystal, experience a large deceleration. The time to decelerate positrons to thermal velocities (the time of thermalisation) is theoretically $\sim 3 \times 10^{-12}$ sec. The mean lifetimes of positrons in various metals is approximately the same for all, and from experimental data is $\sim 1{\cdot}5 \times 10^{-10}$ sec., that is, much longer than the time of thermalisation. The thermalised positrons interact preferentially with free electrons since, with their small energies, they cannot penetrate into the interior of the atoms. The measured width of the spectral lines of the annihilation photons is increased as compared with the width determined from the geometrical parameters of the apparatus; this indicates a large velocity of motion of the free electrons. Study of the profile of the lines or of the angular correlation of the annihilation photons (this is much easier) makes it possible to find out both the mean velocity of the free electrons and the distribution curve of their momenta. Experimental results obtained in this way are in good agreement with the results of the quantum theory of metals.

3. The Quantum Electron Theory of Crystals

Free Electrons in a Metal as Electron Waves

The simplest quantum electron theory of metals, like the classical theory, starts from the supposition of the existence of free electrons. According to wave ideas, a free moving electron is described by a plane monochromatic wave. In a regular undistorted lattice, if $\lambda > 2a$ then such a wave spreads out without scattering and is reflected on the surface of the metal. The stationary states in a metal form a system of plane stationary waves similar to a system of acoustic waves and with a discrete spectrum. Because of the great range of frequencies and the high density of states, this spectrum can be thought of as quasi-continuous and as forming an energy band of finite width (though not greater in width than the magnitude of the internal lattice potential). The electrons are distributed on the energy levels of the crystal in accordance with the Pauli exclusion principle—on each energy level there are not more than two electrons with oppositely directed spins and compensated magnetic moments. Thus the conduction electrons, even at absolute zero, cannot all be placed on the lowest energy level as they can in classical theory. The exclusion principle forces electrons to climb higher up the energy staircase. If in unit volume of a metal there are n free electrons ($n \sim 10^{22}$), then the total number of occupied levels is $\frac{1}{2}n$. We shall now find the position of the highest level of the conduction band occupied by electrons.

Kinetic Energy of the Free Electrons

The number of electron levels, as in the case of the spectrum of free vibrations of a solid (see Chap. 14 §2), is

$$Z = \frac{4\pi V}{3\lambda^3}, \qquad \ldots [8.14]$$

where

$$\lambda = \frac{h}{mv} = \frac{h}{(2m\varepsilon)^{1/2}} \qquad \ldots [8.15]$$

and $\varepsilon = \frac{1}{2}mv^2$ is the kinetic energy of the electrons. Let the volume V of a metal contain N free electrons. Since each level is occupied by two electrons we have

$$Z_{\text{max}} = \tfrac{1}{2}N = \frac{4\pi V}{3\lambda_{\text{min}}^3} = \frac{4\pi V}{3h^3}(2m\varepsilon_{\text{max}})^{3/2}, \qquad \ldots [8.16]$$

whence

$$(2m\varepsilon_{\max})^{3/2} = \frac{N}{V}\frac{3h^3}{8\pi} = \frac{3nh^3}{8\pi},$$

$$\varepsilon_{\max} = \frac{h^2}{2m}\left(\frac{3n}{8\pi}\right)^{2/3} \approx \frac{h^2}{8m}n^{2/3} = \frac{h^2}{8mr_0^2}, \qquad \ldots [8.17]$$

where $r_0^3 = \frac{1}{n}$ is the volume corresponding to one atom.

The maximum energy of the electrons, that is the width of the filled band, does not depend on the total volume of the metal and is determined only by the concentration of free electrons or by the magnitude of the interatomic spacing. If the mass of metal and the total number of free electrons is increased, then as the number of energy levels increases, the distance between neighbouring levels is reduced. If we take it that each atom in a metal contributes one electron, we can determine from formula [*17*] the position of the upper surface of the filled part of the energy band.

$$\varepsilon_{\max} = \frac{40\times10^{-54}}{8\times10^{-27}\times5\times10^{-16}} = 10^{-11}\text{ erg.} = 6\ e\text{V}.$$

Quantum theory gives a much larger value for the maximum energy of the electrons than does the classical theory and leads to the required order of magnitude. An interesting result of the quantum theory is that the large energy of the free electrons is independent of the temperature of the crystal. In fact the energy distribution we have found for the electrons takes the temperature of the crystal to be absolute zero. The electron gas in a metal, as we said above, has an energy corresponding to an effective temperature of the order of several tens of thousands of degrees.

The Energy Distribution Function for Electrons

The distribution function for the occupied energy levels f is essentially different from the Maxwellian distribution. At low temperatures this function obeys the conditions

$$f = 1 \text{ if } \varepsilon \ll \varepsilon_{\max} \text{ and } f = 0 \text{ if } \varepsilon > \varepsilon_{\max} \quad (T = 0), \qquad \ldots [8.18]$$

and this is shown graphically in Fig. 103*a*. At high temperatures ($T \to \infty$) when $kT \to \varepsilon_{\max}$ the specifically quantum effects cannot be observed and we have a Boltzmann distribution (Fig. 103*c*).

$$f \sim e^{-\frac{\varepsilon}{kT}} \quad (T \to \infty). \qquad \ldots [8.19]$$

At intermediate temperatures the edge of the filled band becomes washed out (Fig. 103*b*). The shaded parts above and below the level $\varepsilon_{\max}$ are equal in area corresponding to constancy in the number of particles. As is well known, conditions [*18*] and [*19*] satisfy the Fermi distribution function,

$$f = \frac{1}{e^{(\varepsilon-\varepsilon_{\max})/kT}+1} \qquad \ldots [8.20]$$

which describes the distribution of particles with non-integral spin, that is, the group which includes electrons.

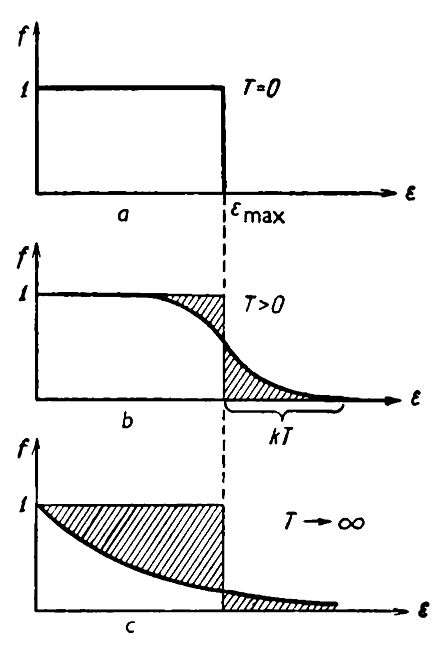


FIG. 103. Transition from Fermi to Boltzmann statistics.

The Specific Heat of the Electron Gas

When there is thermal exchange between the free electrons and the crystal lattice, the change of energy of individual electrons must be equal to the mean thermal energy kT, but only those electrons which lie near to the upper limit of the band can exchange energy with the lattice. Thermal exchange leads to turnover in a thin surface layer of thickness kT from which electrons can be thermally excited to higher unoccupied levels (Fig. 104, compare Fig. 103*b*). Within the filled part of the band, electrons, although they have large kinetic energy, cannot either give their energy to the lattice not receive additional energy since all lower- and higher-lying electron levels are occupied. Thus the 'hot' electron gas cannot heat up the colder lattice. Assuming that the electron levels are distributed with uniform density, we obtain the distance between neighbouring levels as $\Delta\varepsilon = \frac{\varepsilon_{\max}}{\frac{1}{2}N}$. The number of electron levels within the zone of thermal turnover is $\Delta N = \frac{kT}{\Delta\varepsilon} = \frac{1}{2}N\frac{kT}{\varepsilon_{\max}}$. The energy of thermal excitation of electrons in these levels is

$$\Delta W = \Delta N \cdot \tfrac{3}{2}\,kT = \tfrac{3}{4}\,N\frac{(kT)^2}{2\varepsilon_{\max}}. \qquad \ldots [8.21]$$

Hence the electron specific heat is

$$C_e = \tfrac{3}{2}R\frac{kT}{\varepsilon_{\max}} = \gamma T, \qquad \ldots [8.22]$$

where

$$\gamma = \frac{3kR}{2\varepsilon_{\max}}.$$

The specific heat obtained from formula [*22*] is less than that obtained on the ideal gas theory by a factor $kT/\varepsilon_{\max}$, that is by two orders of magnitude for, at room temperature $kT \approx 0{\cdot}03$ eV and $\varepsilon_{\max} \approx 5$-$10$ eV. Consequently at room temperature the electron specific heat

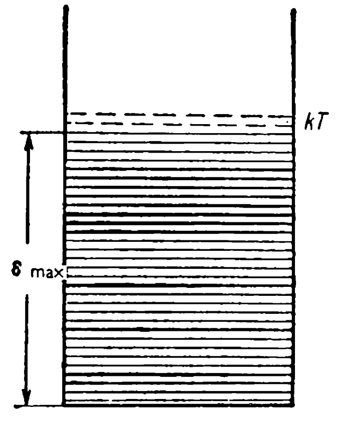


FIG. 104. Calculation of the specific heat of the electron gas.

is only a small part of the specific heat of the lattice C_l. For monovalent metals $C_e \simeq 10^{-2}\, C_l$ which agrees with experiment.

If the number of electron levels in the energy band is half the number of electrons, then all levels are full; there are no free levels under the upper limit of the filled band. The electron specific heat of the crystal in this case is zero. This is the case in dielectrics and semiconductors.

Scattering of Electron Waves in a Crystal

In general, electron waves in a crystal will be scattered. Two types of scattering are possible corresponding to short and long wave-lengths. For a cubic crystal where the length of the edge of the unit cell is a the regions of short and long waves are determined by the relationships

$$\begin{aligned}&(1)\quad \lambda < 2a \quad \text{(short waves)}\\ &(2)\quad \lambda \geqslant 2a \quad \text{(long waves)}\end{aligned} \qquad \ldots [8.23]$$

If $\lambda < 2a$, then the electron waves are scattered by centres which form a discrete three-dimensional periodic structure, and will interfere with each other. If a plane monochromatic wave is propagated in the crystal, an interference pattern arises which is the sum of selective maxima, the position and intensity of which are determined by the symmetry and structure of the crystal and also by the wavelength and direction of the initial wave. This case corresponds to X-ray interference in crystals. The position of the interference maxima is determined by the Bragg-Wulff formula

$$\lambda = 2d \sin\theta, \qquad \ldots [8.24]$$

where d is the interplanar spacing and θ is the angle of incidence. The intensity of the maxima is determined by the formula for the intensity of electron diffraction

$$I_e \approx a/F/^2, \qquad \ldots [8.25]$$

in which the basic quantity is the structure factor $/F/^2$. Formula [*25*] is also analogous to the formula for the intensity of X-ray diffraction I_r, the difference being that in formula [*25*] we must take not the atomic factor f for X-rays but the related function

$$\frac{e^2}{2mv^2}\cdot\frac{1}{\sin^2\theta}(Z-f), \qquad \ldots [8.26]$$

which describes the probability of scattering of electrons by atoms.

If $\lambda \geqslant 2a$, then the electron waves in the crystal do not experience selective reflection, and the crystal can be considered as a continuous medium. This case is analogous to the transmission of visible light through crystalline media: long electron waves passing through a substance experience a scattering analogous to the scattering of light in cloudy media. The intensity of the incoming wave I_0 falls off according to the law,

$$I = I_0 e^{-\mu x}, \qquad \ldots [8.27]$$

where μ is the scattering coefficient and x is the path length of the incoming wave in the medium. Ideally, a regular crystalline lattice is a uniform completely transparent medium for long waves and the scattering coefficient is

$$\mu = 0. \qquad \ldots [8.28]$$

Stresses in the crystalline lattice and non-uniformity in the density of the scattering centres resulting from mechanical deformations, impurities and thermal motion of the atoms, give the crystal a greater or less degree of opacity.

Let us calculate what electron velocity v_{lim} corresponds to the limiting wave length which divides the regions of short and long waves in a crystal with interatomic spacing $a = 3$ Å. From de Broglie's law and condition [*23*] we have

$$v_{\text{lim}} = \frac{h}{2ma} = \frac{6{\cdot}6\times 10^{-27}}{10^{-27}\times 6\times 10^{-8}} \simeq 10^8 \text{ cm/sec.} \qquad \ldots [8.29]$$

This velocity corresponds to an electron kinetic energy of ~ 4 eV.

The free electrons of a metal which occupy the highest energy levels have a velocity $v \approx v_{\text{lim}}$. The waves corresponding to them are short and the motion of such electrons in the crystal will be accompanied by phenomena of selective reflection.

In an electric field, the free electrons have also a directed motion with velocity u along the lines of force of the field, corresponding to an electric current in the metal. In a field of strength

$$E = 1 \text{ volt/cm} \approx 3\times 10^{-3} \text{ c.g.s. units}$$

$$u = \frac{eE}{2mv}\,l = \frac{5\times 10^{-10}\times 3\times 10^{-3}}{2\times 10^{-27}\times 10^{8}}\,l \simeq 10^7\, l \text{ cm/sec.} \qquad \ldots [8.30]$$

For any physically credible value of the electron mean free paths in the metal ($l \sim 10^{-8}$ cm on the classical theory or $l \sim 10^{-6}$ cm on the quantum theory) the velocity of electrons in the electric field

$$u \ll v_{\text{lim}}. \qquad \ldots [8.31]$$

This velocity corresponds to the lengths of electron waves which are scattered by irregularities and lattice distortions.

The Electrical Conductivity of Metals

The dual nature of microscopic particles which have corpuscular and wave properties means that we can look at the same phenomenon from different points of view. Consider the scattering of a beam of corpuscles which are passing through a medium

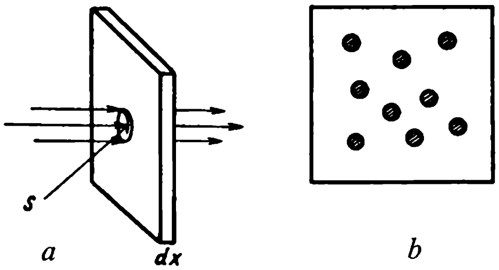


FIG. 105. The law of scattering.

which contains scattering centres. Let n be the density of scattering centres in unit volume. Consider a thin layer of the medium of thickness dx (Fig. 105*a*) and let S be the area of the cross-section of the beam of corpuscles. In the irradiated volume there are $dn = nSdx$ particles. If each centre is a target with a scattering radius r (Fig.

105*b*) then the total cross-sectional area of all centres is $dS = n\pi r^2 S dx$. The mean free path l of the particles is equal to the thickness of a layer for which the total area of all scattering cross-sections is equal to the total area of the cross-section. This condition gives

$$l = \frac{1}{n\pi r^2}. \qquad \ldots [8.32]$$

The proportional decrease of the number of particles after passing through a layer of thickness dx will be

$$-\frac{dI}{I} = n\pi r^2 dx.$$

And from this we obtain the scattering law for the particles.

$$I = I_0 e^{-n\pi r^2 x} = I_0 e^{-\frac{x}{l}}. \qquad \ldots [8.33]$$

Comparing [*33*] with [*27*] we obtain an expression for the scattering coefficient μ in terms of the mean free path of the particles l and the effective cross-section of the scattering centres

$$\mu = \frac{1}{l} = n\pi r^2. \qquad \ldots [8.34]$$

Putting the scattering coefficient into the formula for the specific resistivity of metals [*7*] we have

$$\rho = \frac{2mv}{ne^2}\frac{1}{l} = \frac{2mv}{ne^2}\mu. \qquad \ldots [8.35]$$

For an ideal regular crystal without impurities, scattering is caused only by fluctuations in the density of particles produced by thermal vibrations. The radius of the scattering cross-section is determined by the amplitude of thermal vibrations of the particles ξ,

$$\mu_2 = n\pi\overline{\xi}^2.$$

Now the mean value of the elastic energy of an oscillator is

$$\tfrac{1}{2} f\overline{\xi}^2 = \tfrac{1}{2} kT,$$

where f is the coefficient of quasi-elastic force. So

$$\mu_T = \frac{n\pi kT}{f} = \frac{n\pi k}{aE} T, \qquad \ldots [8.36]$$

where E is Young's modulus and a is the lattice constant.

Formulae [*35*] and [*36*] show that the specific resistivity of a pure metal is a linear function of temperature. This is in agreement with experiment. In a metal which has been distorted as a result of mechanical deformation and in which there are also impurities present the scattering coefficient for electron waves is made up of three parts

$$\mu = \mu_{\text{def}} + \mu_{\text{imp}} + \mu_T. \qquad \ldots [8.37]$$

At low temperatures the scattering of electrons is determined by the first two terms in [37]; these represent the residual electrical resistivity of metals.

4. The Band Theory of Crystals

Energy Bands

For a freely moving electron in a field of constant potential the kinetic energy is constant and proportional to the square of the modulus of the wave vector:

$$\varepsilon = \tfrac{1}{2} mv^2 = \frac{h^2}{2m} k^2. \qquad \ldots [8.38]$$

Graphically, the dependence of ε on k is represented by a parabola (see Fig. 1, page 11). In a crystal lattice, the potential is a periodic function of distance and is only constant on the average. For the motion of electrons in a periodic field, if $|k| > 2a$, we have selective reflection of the waves, and to this there corresponds energy gaps in the curve of $\varepsilon = \varepsilon(k)$ (see Fig. 22, page 45). The quasi-continuous electron spectrum of a crystal is divided into bands of allowed and forbidden energy values. The directions of the wave vectors $k_1, k_2, \ldots$ corresponding to forbidden values of the energy, are determined by the conditions of the Bragg formula:

$$k = \frac{1}{\lambda} = \frac{1}{2d \sin \theta} = \frac{H}{2 \sin \theta}, \qquad \ldots [8.39]$$

where H is the modulus of a vector of the reciprocal lattice. The magnitudes of the energy discontinuities $\Delta_{\varepsilon_1}, \Delta_{\varepsilon_2} \ldots$ are proportional to the squares of the moduli of the structural amplitudes $|F_i|^2$ for the corresponding planes of the crystal.

The geometry of the energy discontinuities is conveniently shown in wave vector space (k-space) which, as formula [39] shows, is closely connected with reciprocal lattice space: consider in the crystal, any system of planes with Miller indexes (hkl) and described by a vector of the reciprocal lattice $\boldsymbol{H}$, the modulus of which is $H = 1/d$ (Fig. 106). The geometrical position of the vectors $\boldsymbol{k}$ which satisfy the Bragg Law is represented in k-space by two parallel planes AA and BB situated at distances

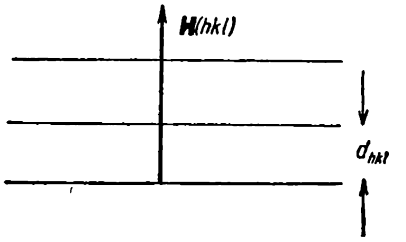


FIG. 106. A family of planes (h, k, l) and the vector H of the reciprocal lattice.

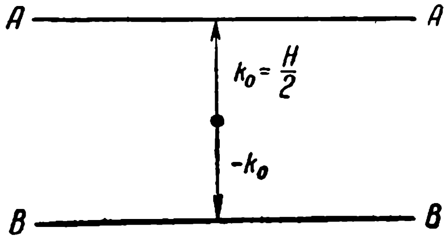
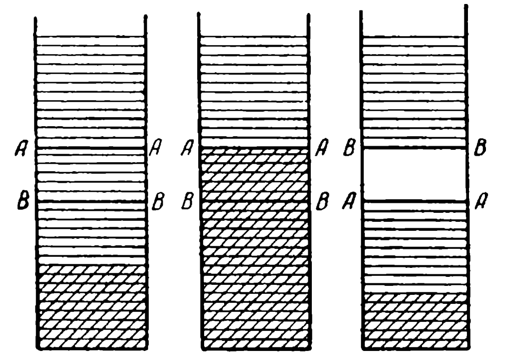


FIG. 107. The distance between planes in k-space.

$\pm k_0 = \frac{H}{2} = \frac{1}{2d}$ from the origin (Fig. 107). The direction of the normal to these planes is parallel to the vector $\boldsymbol{H}$. If the crystal is symmetrical, then the system of equivalent planes $\{hkl\}$ forms a solid figure. Each figure $\{hkl\}$ corresponds to a similar one (in k-space) for the energy discontinuities (a zone) and the dimensions of this zone are determined by the reciprocals of the interplanar spacing $1/d$. In a crystal, when the indices $\{hkl\}$ of the planes are increased, the interplanar spacing is, in general, reduced and each successive figure is enclosed within the previous one. In k-space when the indices are increased the zone dimensions increase and each succeeding zone encloses the previous one. Corresponding to the conditions for X-ray extinction whereby the structural amplitudes are divided into two groups, $F = 0$ and $F \neq 0$, the first zone for a simple cubic lattice is a cube while, for a b.c.c. lattice it is a rhombic dodecahedron and for an f.c.c. lattice, an octahedron intersected by a cube (a cubo-octahedron).

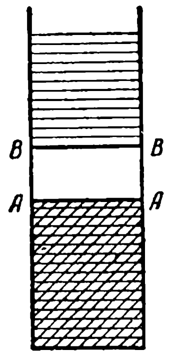
The Electron Spectrum of a Metal and an Insulator

The successive zones in an arbitrary crystal may or may not intersect each other. In the individual zones there are also two possibilities; the levels of the zone may be either completely filled or incompletely filled with pairs of electrons. The different possible types of energy spectra of crystals are shown in Figs. 108 and 109. Here, in all the diagrams, the line AA marks the upper limit of the first zone and the line BB the bottom of the second zone; the horizontal lines are possible electron levels and the diagonal lines show the regions of filled levels. Figs. 108a and b correspond to an overlap of zones and Figs. 108c and 109 correspond to energy gaps. In Figs. 108a and c there is partial filling of the first zone and in Figs. 108b and 109 these zones are completely full. The first three types of energy spectra correspond to crystals with metallic properties since there are free levels immediately above the filled levels. If such a crystal is brought into an electric field, the electrons in the upper strip of the filled region can be accelerated by the electric field and move to higher levels. As a result of inelastic collisions these electrons will move back to lower free levels and the difference in energy in these transitions appears in the form of Joule heat. The case of Fig. 109 is that of an insulator: the first energy zone is here completely filled and the second zone which is separated from the first by the energy discontinuity $\Delta E \gg kT$ is completely empty. In a crystal with such an energy spectrum the electric field



metal a metal b metal c

FIG. 108. The zone scheme for a metal.



insulator

FIG. 109. The zone scheme for an insulator.

cannot accelerate the electrons. Such a crystal will have the properties of an insulator although in it, as in a metal, there is a large collective electron pool.

The Electron Spectrum of Semiconductors

The spectrum of a semiconductor is similar to that of an insulator: the first energy zone is completely filled and the second is separated from it and is completely empty. However, the magnitude of the energy discontinuity ΔE between the zones is small, much smaller than in an insulator but larger than kT (Fig. 110*a*). At ordinary temperatures few electrons can be moved by thermal activation from the first zone to the second. When however, the crystal is heated, a number of electrons can make the transition to the second zone and allow the crystal to conduct an electric current.

When Δn electrons make the transition into the upper, previously empty, band the same number of electron levels in the lower, formerly completely full, band are freed, and these can be thought of as 'holes' in the energy spectrum of the crystal. The transport of electricity in a semiconductor takes place by means of electrons in the conduction band and also by 'holes' in the filled band. In both cases, charge is really transported by electrons, but it is convenient to consider conductivity in the incompletely filled band as a movement of holes corresponding to a movement of particles with an effective mass and with charge equal to the charge on the electron but with the opposite sign. Thus the mechanism of conductivity in a semiconductor involves two particles. In the conduction band there appear Δn^- negative electrons and in the filled band Δn^+ electrons with an effectively positive charge. In the case of thermal excitation of the electrons from the lower into the upper band, the number of charge carriers of both types is the same; $\Delta n^+ = \Delta n^-$. In the general case this is not necessary and there may be charge carriers either of the first or of the second type, or there may be an excess of the first or the second type. Thus we distinguish semiconductors of the Δn^+ type and the Δn^- type. Experimentally, the distinction between the two types of semiconductors is made by the sign of the Hall constant, that is by experiments on the passage of current in a semiconductor in a magnetic field. Because of the smallness of the electron and hole concentration, the electrical conductivity is mainly determined by the number of charge carriers, which increases with temperature. The electrical resistivity of semiconductors, contrary to that of metals, thus falls on heating.

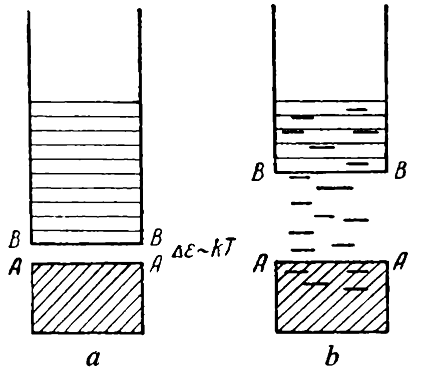


FIG. 110. The zone scheme for a semiconductor: (*a*) with a narrow gap, (*b*) with impurity levels.

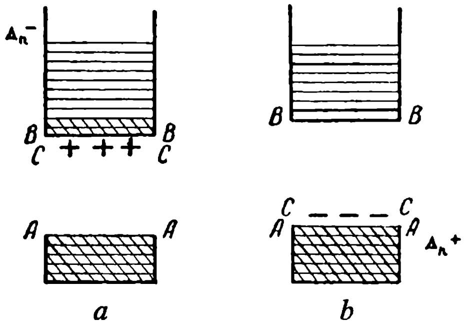


FIG. 111. The zone for a semiconductor (*a*) electron type, (*b*) hole type.

'Holes' in the energy spectrum of a crystal must not be confused with atomic 'vacancies' which are empty sites in the atomic crystalline lattice structure and which are important in diffusion and self-diffusion of atoms in crystals. There is, however, a connection between vacancies and holes in that damage and disturbances in the atomic lattice, including such disturbances as vacancies, atoms displaced by deformation, and impurities, give rise to distortions and defects in the energy spectrum of the crystal. The simplest view of the distorted spectrum is that of a system of local energy levels superposed on the basic crystal spectrum. These local levels are indicated by heavy lines in Fig. 110*b* and those which fall within the energy gap form intermediate energy steps on which electrons can be placed or from which electrons can be removed.

This scheme of local levels enables us to understand the great sensitivity of the electrical and optical properties of semiconductors to lattice distortions and impurities, and also enables us to explain the properties of the impurities. Two types of impurity can be distinguished by their chemical nature; those which are more and those which are less electronegative than the semiconductor itself for on this property depends the outcome of the struggle for an electron between atoms of the impurity and atoms of the semiconductor. Impurities can be donors or alternatively, acceptors of electrons. Sometimes the same impurities can be either donors or acceptors.

Consider the atoms of a donor impurity represented by the local levels *CC* (Fig. 111*a*) which lies close to the bottom of the conduction band within the limits of the strip for which thermal activation is effective. Electrons from the donor atoms go over into the conduction band and as a result produce conductivity of the Δn^- or *n*-type. This type of conductivity takes place in zinc oxide, which when heated in a reducing atmosphere loses a small fraction of oxygen atoms from its lattice. Neutral zinc atoms are thus left and distributed throughout the lattice by diffusion. When this zinc oxide is heated it develops electron conductivity because the zinc atoms act as donors transferring some of their electrons to the conduction band.

Now consider impurity atoms, this time acceptors, represented by the local levels *CC* (Fig. 111*b*) which lie close to the upper boundary of the filled zone and within the limits of the strip for which thermal activation is effective. Some of the electrons from the filled band will be captured by atoms of the impurity, leading to the appearance of a number of holes and to the development of conductivity of the Δn^+ type. This type of conductivity is found in copper oxide (CuO), which when heated in an oxidising atmosphere (in oxygen) absorbs a certain amount of oxygen. Oxygen atoms which have diffused into the lattice are acceptors, and when a specimen of copper oxide which has been treated in this way, is heated they capture some of the electrons from the filled band and become negative ions. As a result of this we have hole conductivity of the Δn^+ or *p*-type.

In semiconducting crystals of elements of the IV B subgroup (silicon, germanium) atoms of neighbouring elements from the III B or V B subgroups can replace a small number of atoms of the basic element and form a substitutional solid solution. Thus, silicon forms such solutions with boron (III B subgroup) and also with phosphorus (V B subgroup). Atoms of elements of B subgroups lying to the left of the IV B subgroup have insufficient valence electrons and are acceptors while atoms of elements lying to the right have an excess of valence electrons and are donors.

The Electron Spectrum of an Ionic Crystal

The electron formulae for the atoms of lithium and fluorine are respectively $^3\text{Li}\ 1s^2\,2s^1$ and $^9\text{F}\ 1s^2\,2s^2\,2p^5$. Valence electrons of the $2p^5$ level of the fluorine

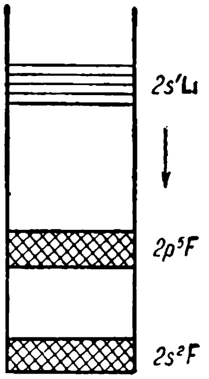


FIG. 112. Zone scheme in the LiF crystal.

atoms are situated below the 2*s* level of lithium. This corresponds to a large electronegativity of the fluorine atom as compared with the lithium atom. Thus the 2*s* electron of lithium transfers to the 2*p* level of fluorine, and results in the formation of the ions Li^+ and F^-. As a result, the 2*s* band of lithium is completely emptied and the 2*p* band of fluorine completely filled (Fig. 112). Since these bands do not intersect, the crystal of lithium fluoride does not exhibit electron conductivity and is an insulator. At high temperatures, however, conductivity of the ionic type is developed in it.

The Electron Spectrum of Metals

The diagram of the splitting of sodium levels as a function of the distance between atoms is shown in Fig. 113.

The electron formula for the sodium atom is $1s^2\ 2s^2\ 2p^6\ 3s^1$; the 3*p* subgroup is not occupied. Now the 3*p* level begins to split at larger separations than the 3*s* level since the 3*p* state corresponds to a greater distance of the electrons from the nucleus than does the 3*s* state. As a result of the splitting, the 3*s* and 3*p* zones intersect and the crystal of sodium exhibits metallic properties. Fig. 114 shows the splitting of the electron levels in copper (^{29}Cu $3d^{10}\ 4s^1$), where the 4*s* and 4*p* levels are broadened earlier than the 3*d*. When $r = r_0$ the 3*d* band is much narrower than the 4*s* and 4*p* bands. In copper, the 3*d* band is completely filled by 10 electrons per atom, so in this band there is a high density of levels. In the 4*s*-*p* band there is one electron. The relative positions of the *s*-*p* and *d* bands in copper and their filling are shown schematically in Fig. 115. The nearest neighbour to copper is nickel, which has one electron less than copper but the same crystal structure. The zone scheme for nickel is shown in Fig. 116. The removal of the one electron could have completely emptied the 4*s*-*p* band, but in fact, some of the electrons come out of the 3*d* band. The upper limit of the filled levels in both bands will be at the same height. The 3*d* band of nickel thus has 'holes' and from the magnetic data it turns out that these 'holes' correspond on

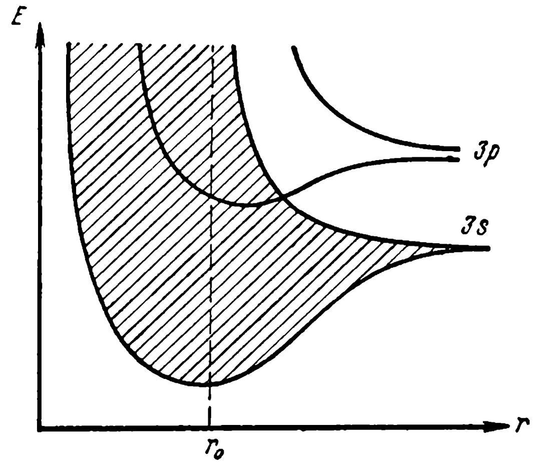


FIG. 113. Splitting of the sodium levels as a function of inter-atomic spacing; r_0 = the distance between atoms in the sodium crystal.

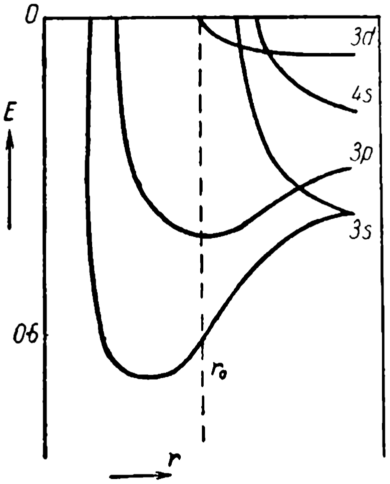


FIG. 114. Splitting of the energy levels of copper.

the average to 0·6 electrons per atom. Thus the 4*s*-*p* band of nickel is filled up to a level corresponding to 0·6 electrons per atom. Comparatively small differences in the electron spectra of nickel and copper are connected with fundamental differences in their physical properties. These differences are basically as follows:

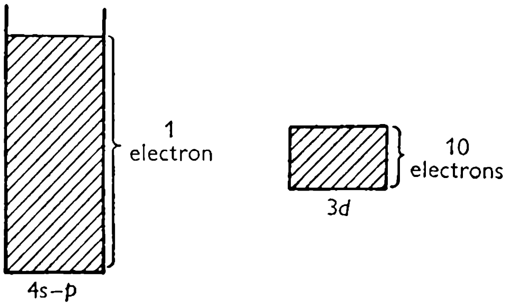


FIG. 115. Energy bands in copper.

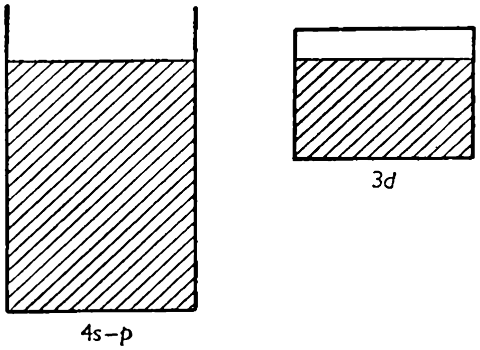


FIG 116. Energy bands in nickel.

(1) The copper crystal is diamagnetic while the nickel crystal is ferromagnetic. Ferromagnetism of nickel, cobalt and iron is due to incomplete filling of the 3*d* band and the positive sign of the exchange integral (see chapter 10 §4). The uncompensated magnetic moment of the 3*d* electron spins is the basis of their ferromagnetism. In copper on the other hand, the 3*d* band is completely full and the magnetic moments of 3*d* electrons are fully compensated, so copper is diamagnetic.

(2) The electrical resistivity of nickel is 14 times larger than that of copper. This enhancement of the electrical resistivity of nickel and the other transition metals can be explained by the reduction in the mean free path of the electrons which results from the possibility that conduction electrons can be scattered not only within the conduction band but also from one band to the other.

(3) At high temperatures, the specific heat of nickel and other transition metals is in excess of the value to be expected from Dulong and Petit's Law although the lattice specific heat in these metals is the same as in metals of the copper group (chapter 14 §1). The difference comes from the electron specific heat, for in nickel and other transition metals with incompletely filled *d* bands, thermal excitation of the *d* electrons is possible and the high density of states in the *d* band means that a large number of electrons can be excited.

From the energy band scheme we can understand some of the colour changes commonly observed in alloy systems. In the copper-zinc alloys, for instance, the colour changes from the red ('copper') colour of pure copper through the yellow ('brassy') of the alloys of intermediate concentration to the greyish-white metallic colour of zinc. Increases in the concentration of zinc are accompanied by an increase in the concentration of conduction electrons since copper is monovalent while zinc is divalent. Fig. 117 shows the changes of the degree of filling of the 4*s*-*p* band as a function of zinc concentration. Now we may assume that absorption of light in copper takes place as a result of electron transitions from the *d* band to empty levels in the *s*-*p* band. The energy of the light quantum necessary is indicated by the small arrow and corresponds to the yellow part of the spectrum. Thus all quanta with smaller wavelength can excite this transition and be absorbed. In Fig. 118 the part of the incident white spectrum which is absorbed is shown by the shading. The part of the spectrum with longer wavelengths will not be absorbed (Fig. 118*a*); it is reflected and scattered at the copper surface giving it its reddish shade. As zinc is added the level to which the *s-p* band is filled rises and the limiting wavelength for absorption moves towards the blue end of the spectrum (Fig. 118*b*). When the composition is 70% copper and 30% zinc the quanta which are absorbed in the alloy correspond to the blue region of the spectrum and the colour of the alloy becomes yellowish. Eventually at the zinc end, the absorption moves into the ultra-violet (Fig. 118*b*) and all the visible part of the spectrum is reflected and scattered. Thus metallic zinc, like many other metals, is colourless.

5. The Theory of Electron Compounds

In alloys of monovalent copper, silver and gold with elements of the B subgroups which have higher valency, there appear a series of typical phases called electron compounds. As in many metallic compounds, the electron compounds have more or less wide solubility ranges (chapter 6 §3).

In distinction to ionic or covalent compounds where the formation of the uniform type of structure requires a unique stoichiometric composition, in the electron compounds this is not necessary. Phases which are uniform in structure appear at certain definite concentrations of the electron gas; atomic concentrations may vary. To determine the electron concentration we need to know how many electrons each atom of one or the other metal contributes to the common pool. For elements of the B subgroups we may take these electrons to be the electrons of the *s* and *p* levels, the number of which agrees with the basic valency of the corresponding elements. In the transition metals (iron, cobalt, nickel) which have unfilled *d* levels the calculation of the electron concentration requires the contribution of the corresponding atoms to be taken as zero. This is in fundamental disagreement with the chemical valency which these elements have in compounds with atoms of the metalloids. Physically, this behaviour of the atoms of the transition metals can be explained by their desire to fill the free levels in the *d* band. Table 47 gives values of the number of electrons contributed by various atoms to the general pool of electrons and which determine the formation of electron compounds.

In the compounds we are considering, an increase in the atomic concentration of the second component is accompanied by an increase in the concentration of

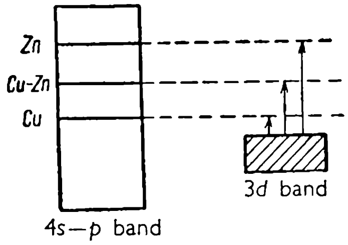


FIG. 117. Filling of the 4 (*s*-*p*) band of copper as a function of the amount of zinc in the alloy.

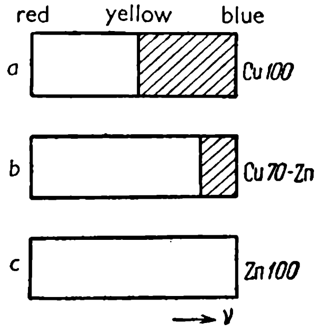


FIG. 118. Explanation of the colours of the copper-zinc (brass) alloys.

Table 47

Electron Contributions from Metals in Electron Compounds

Subgroup	Metals	Electron Contribution	Chemical Valency
VIIIA	Fe, Co, Ni, Ru, Ph, Pd, Pt, Ir, Os	0	
IB	Cu, Ag, Au	1	1
IIB	Be, Mg, Zn, Cd, Hg	2	2
IIIB	Al, Ga, In	3	3
IVB	Si, Ge, Sn	4	4
VB	P, As, Sb	5	5

electrons. In k-space, the filled electron levels are represented by points which lie within a sphere of radius $k_{\max}$ corresponding to [*17*] $\varepsilon_{\max} = \frac{h^2}{2m}\left(\frac{3n}{8\pi}\right)^{2/3}$ (the Fermi Sphere). Increasing the electron concentration leads to an increase in the radius of the sphere of filled levels until it intersects the polyhedral-shaped zones formed by planes of energy discontinuity. When this happens, some electrons have to make the transition into the following energy band because the preceding band is full. However because of the energy discontinuity ΔE at the boundary of the zone (Fig. 119) the transition is accompanied by a discontinuous increase in kinetic energy of the electrons, leading to a lowering of the stability of the phase. Consider now a phase with another structure; if the capacity of the first zone of this phase is larger so that all the conduction electrons in the alloy can move into it, then a transition to that structure is probable. Let us calculate the capacity of the first energy zones for various electron compounds.

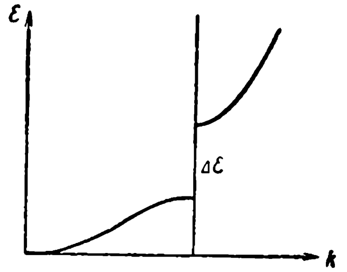


Fig. 119. Energy discontinuity at a zone edge.

Face-centred Cubic Structure (α-phase)

The structural amplitude has the form

$$F(h\,k\,l) = 1+\cos\pi\,(h+k)+\cos\pi\,(k+l)+\cos\pi\,(l+h). \quad \ldots [8.40]$$

From which we see that

$$F(h\,k\,l) \neq 0 \quad \ldots [8.41]$$

only in the case when h, k, and l are not mixed, that is, when all three indices are simultaneously even or uneven. The successive reflexions for the α-phase have the following indices: 111, 200, 220, ... Considering these indices as co-ordinates of points in the reciprocal lattice, we arrive at a b.c.c. structure. To obtain the zones we have to draw vectors of the reciprocal lattice H from any lattice point in such a way that $k_n = \frac{1}{2}d = H/2$, divide them in two and draw perpendicular planes. The first zone is a combination of an octahedron {111} and a cube {200}; that is a cubo-octahedron. The volume of the first zone referred to one lattice site is

$$v_z = \frac{1}{2}\,\frac{8}{a^3} = \frac{4}{a^3}. \quad \ldots [8.42]$$

The radius of a sphere described in the first zone and touching the octahedral faces will be

$$R = \tfrac{1}{2}\frac{\sqrt{3}}{a},$$

and the volume of this sphere is

$$v_{\text{sph}} = \tfrac{4}{3}\pi R^3 = \frac{\pi\sqrt{3}}{2a^3}. \qquad \ldots [8.43]$$

The number of electron states in the first zone is equal to the volume of the sphere. The concentration of conduction electrons n_e in the alloy, relative to the number of atoms n_a, is equal to twice the ratio of the volumes

$$\frac{n_e}{n_a} = 2\frac{v_{\text{sph}}}{v_z} = 2\frac{\pi\sqrt{3}}{8} = 1{\cdot}36, \qquad \ldots [8.44]$$

since each electron level is occupied by two electrons. This ratio agrees well with experiment and gives the limit of solubility of the α-solid solution.

Body-Centred Cubic Structure (β-phase)

The structural amplitude has the form

$$F(h\,k\,l) = 1+\cos\pi\,(h+k+l), \qquad \ldots [8.45]$$

$$F(h\,k\,l) \neq 0 \quad \text{when} \quad h+k+l = 2n. \qquad \ldots [8.46]$$

The succession of indices (h, k, l) giving coherent reflexions is: 110, 200, 211, 220, 310 which, in reciprocal space, form a face-centred lattice. The first zone in the face-centred cube is bounded by planes of a rhombic dodecahedron {110}. The volume of the first zone, referred to one lattice site is

$$v_z = \tfrac{1}{4}\frac{8}{a^3} = \frac{2}{a^3}. \qquad \ldots [8.47]$$

The radius of the inscribed sphere is

$$R = \frac{2}{a}\frac{\sqrt{2}}{4} = \frac{1}{a\sqrt{2}},$$

and the volume of the sphere is

$$v_{\text{sph}} = \tfrac{4}{3}\pi R^3 = \frac{4\pi}{3}\cdot\frac{1}{a^2 2\sqrt{2}} = \frac{2\pi}{a^3 3\sqrt{2}}. \qquad \ldots [8.48]$$

Hence the electron concentration is

$$\frac{n_e}{n_a} = 2\frac{v_{\text{sph}}}{v_z} = \frac{\pi\sqrt{2}}{3} = 1{\cdot}48, \qquad \ldots [8.49]$$

which is close to the value 3/2 as found by experiment.

The γ-phase has a cubic lattice with 52 atoms in the unit cell, very similar to the super-lattice of the β-phase and obtained from it by distorting it slightly.

Structural Amplitudes F (*hkl*) for the γ-Phase from X-Ray Data

h k l	110	200	211	220	310	220	321	400	330	411
F(*hkl*)	0·32	0	0·32	0	0·32	2·68	1·05	0	8·85	5·63

The highest reflectivity is that of the {330} and {411} planes, which form a zone figure with 36 faces, closely resembling a sphere. Calculation of the volume of the zone gives $v_z = 90$ electrons while the volume of the sphere is $v_{\text{sph}} = 80$ electrons. By experiment $\frac{n_e}{n_a} = \frac{21}{13} = 1{\cdot}62$. With 52 atoms in the unit cell this corresponds to a limiting concentration of 84 electrons.

Hexagonal Close-Packing (ε-phase)

Calculation of the capacity of the first zone gives the limiting concentration as $n_e/n_a = 1{\cdot}74$ while experiment gives $n_e/n_a = \frac{7}{4} = 1{\cdot}75$.

Chapter 9

ELECTRICAL PROPERTIES OF CRYSTALS

1. Classification of Solids by their Electrical Properties

Classification by the Type of Conductivity

Since the electrons play such an important part in determining physical and chemical properties, it is appropriate to divide all solids into two basic groups, *metals* and *non-metals*, the principal difference between which is the form of their electron energy spectra. We consider as non-metals those substances with a finite value for the activation energy of the valence electrons, $\Delta E > 0$, and as metals those substances for which $\Delta E = 0$.

Metals are good electrical conductors and are characterised by the smallest values of the specific resistivity. The electrical conductivity of a metal falls as the temperature rises. In non-metals the valence band is full and the conduction band is completely empty, so at low temperatures the substance behaves as an insulator. As the temperature is raised, thermal activation takes place; some of the valence electrons make the transition into the conduction band and produce electron conductivity. The electrical conductivity increases rapidly with temperature according to an exponential law. Such a substance is called a semiconductor and its electrical conductivity, produced by temperature activation, is called *intrinsic* conductivity.

The electrical conductivity of a non-metal is extremely sensitive to impurities. If there are no impurities and if the width of the forbidden band ΔE is sufficiently large compared with kT, the substance behaves as an insulator. Addition of special impurities or distortion of the atomic structure may cause the formation of local energy levels which partially fill the forbidden band. The distribution and number of these levels is determined by the type of distortion, by the nature of the impurities, and by their concentration. In impurity semiconductors various processes may take place: (*a*) some of the valence electrons may be captured by the impurity levels, (*b*) electrons may be transferred from these levels into the conduction band, (*c*) both processes may take place together. In all these cases electron conductivity arises. If the first process is predominant then the mechanism of conductivity is called '*hole*' type (p-type) and if the second is predominant it is '*electronic*' (n-type). The electrical conductivity of a semiconductor arising from impurities is called *impurity* conductivity or *extrinsic* conductivity.

In non-metals, where there is no electronic conductivity there may be ionic conductivity which is a possible method of charge transfer in heteropolar substances at high temperatures, when the velocity of diffusion is sufficiently large. The electrical conductivity in this case, just as in semiconductors, increases exponentially with temperature. In solid electrolytes and also in solutions, ionic conductivity is accompanied by electrolysis. Electronic and ionic conductors are called respectively conductors of types I and II.

The electrical conductivity of solids depends on their physical and chemical nature (composition, atomic and electronic structure) and the conditions of the experiment (temperature etc.) and varies within wide limits (Table 48).

Table 48

Specific Resistivity of Some Solids

Insulators and Semiconductors	ρ (ohm.cm)	Metals	ρ (ohm.cm)
Paraffin Wax	3×10^{18}	Bismuth	$1{\cdot}2 \times 10^{-4}$
Sulphur (70° C)	4×10^{15}	Tellurium	2×10^{-5}
Quartz	$1{\cdot}2 \times 10^{14}$	Lead	2×10^{-5}
Diamond	$10^{12}-10^{13}$	Iron	$1{\cdot}2 \times 10^{-5}$
Silicon	$\sim 10^2$	Aluminium	$3{\cdot}2 \times 10^{-6}$
Germanium	~ 10	Silver	$1{\cdot}7 \times 10^{-6}$
Graphite	3×10^{-3}	Superconductors	0

The ratio between the electrical resistivity of the best insulators and the best conductors attains the astronomical figure of $\sim 10^{24}$. Some metals at low temperatures go over into the *superconducting* state and their resistivity is zero. The ratio just mentioned is in this case infinite.

Classification of Dielectrics

At sufficiently low temperatures and in fields which are weak relative to the minimum electrical strength of the substance, non-metals behave as insulators or dielectrics. When these are brought into an electrical field a current does not pass but the field produces a series of essential changes in their atomic and electrical structure and their electrical state. These changes are called electrical polarisation. The polarisation P is measured by the sum of the moments in unit volume of the body

$$\boldsymbol{P} = \sum \boldsymbol{p}_i, \qquad \ldots [9.1]$$

where the $\boldsymbol{p}_i$ are the elementary electrical moments of the particles of the substance.

The dielectric properties are described by the dielectric constant ε (or the susceptibility η), which is connected with the induction $\boldsymbol{D}$, the field $\boldsymbol{E}$, and the polarisation $\boldsymbol{P}$ by the relations,

$$\boldsymbol{D} = \varepsilon \boldsymbol{E} = \boldsymbol{E} + 4\pi \boldsymbol{P}, \qquad \ldots [9.2]$$

from which we get

$$\varepsilon = 1 + 4\pi P/E = 1 + 4\pi\eta. \qquad \ldots [9.3]$$

The dielectric constant is determined by the change of capacity of a condenser when it is filled with dielectric. It is often necessary to produce substances which have large values of the dielectric constant and, as equation [2] shows, this depends on the polarisability of the body.

In molecules and crystals, which contain heteropolar atoms, there are constant elementary electric moments, the magnitudes of which depend on the differences of electronegativity and on the interionic distances. The magnitude of the resultant moment is determined by the symmetry and structure of the crystal, which can be considered as an ordered system or systems of dipoles. The resultant moment is equal to the geometric sum of the moments of the individual sub-lattices of dipoles which can be picked out from the crystal structure. In crystals with a centre of symmetry the resultant moment is zero although the elementary moments may be large (for example, ionic crystals of the Na^+Cl^- type). In crystals without a centre of symmetry the resultant moment may differ from zero. Such crystals, called *electrets*, have constant polarisation and are analogous to permanent magnets. The electret state is more difficult to discover than that of spontaneous magnetization since the electrically charged particles and ions from the surrounding medium are attracted by the electret and compensate the polarising charges on its surface.

> 'Electrets' was initially the name given to some amorphous substances in which constant polarisation was produced artificially by special treatment. In amorphous substances which contained polar groups (for example, some resins) the resultant moment was zero because of the randomness of their structure. An oriented distribution of the molecules can be produced if the resin is melted in an electric field and the polarised state can be made permanent by cooling the resin in the electric field. Artificial amorphous electrets were produced by Aguchi (1925) by means of special thermal and electric treatments.

When the temperature of the crystal is changed, the interatomic spacings and valence angles are changed as a result of thermal expansion. In a crystalline electret this leads to a change in the magnitude of the polarisation. This phenomenon is called pyroelectricity, and the corresponding crystals are called *pyroelectrics*. Pyroelectricity can exist in all electrets. The state of polarisation of crystals is changed as a result of mechanical deformation (piezoelectricity). The piezoelectric effect can also be excited in some crystals which do not have constant polarisation and thus the group of *piezoelectrics* is much wider than the group of electrets. To the group of electrets there also belong crystals in which the direction of polarisation changes easily in weak electric fields. Such substances are analogous to soft ferromagnetic materials and are called *ferroelectrics*. The ferroelectric state is easily destroyed by changes of temperature and the ferroelectric then passes into an unpolarised paraelectric state. *Antiferroelectrics*, analogous to antiferromagnetics, is the name given to substances in which a change of temperature gives rise spontaneously to an ordered distribution of the elementary dipole moments with full compensation of the resultant moments of the individual sub-lattices of the crystal.

2. The Dielectric Constant and Polarisability

The Polarisability of Atoms and Ions

In the free atom or ion the centre of gravity of the electron cloud, averaged over time, coincides with the nucleus. The electric moment of the atom is thus zero. An external electric field E brings about a movement of the electron cloud relative to the nucleus and induces an electric moment p_i in the atom. The magnitude of this moment depends on the strength of the field and on the deformability of the electron cloud, the latter being determined by the relationship between the external field and the internal atomic field. The field acting on the electron in a hydrogen atom is

$$E_H = e/a_0^2 \approx 6\times 10^9 \text{ V/cm}$$

The strength of this field is much larger than the strength of fields which we can produce; and which can be considered weak in comparison with atomic fields. The electric moment of an atom in a weak field can, with sufficient accuracy, be taken as proportional to the field strength

$$\boldsymbol{p} = \alpha \boldsymbol{E} \qquad \ldots [9.4]$$

The coefficient α is called the polarisability of the atom and has the dimensions of a volume.

The atomic polarisability can be found for a given model of an atom. In the classical model the bound electrons are considered as harmonic oscillators, the natural frequency ω of which is determined by formula [*1.130*] where $f = k_q$ is the coefficient of quasi-elastic force of the qth electron. In an electric field E, the electron is displaced —from its position of equilibrium—by a distance x_q, determined by the expression

$$eE = k_q x_q = m\omega_q^2 x_q. \qquad \ldots [9.5]$$

From [*4*] and [*5*] we obtain the contribution of the qth electron to the atomic polarisability:

$$\alpha_q = p_q/E = ex_q/E = e^2/k_q = e^2/m\omega_q^2 \qquad \ldots [9.6]$$

The atomic polarisability is the sum of the polarisability of all electrons in the atom:

$$\alpha = \sum \alpha_q = (e^2/m) \sum \omega_q{}^{-2} \qquad \ldots [9.7]$$

In this sum the most important terms are those for the valence electrons which are situated in the most easily deformed part of the electron cloud. The coefficient of quasi-electric binding for a hydrogen atom is $k = e^2/a_0^3$ (see Problem 1 at the end of the chapter). The polarisability of the hydrogen atom is

$$\alpha_H = a_0^3 = 0{\cdot}15\times 10^{-24} \text{ cm}^3,$$

and this is proportional to the atomic volume and agrees in order of magnitude with experiment (see Table 52, page 252).

In a periodic electric field of frequency ω, the bound electrons execute forced oscillations, the amplitude of which depends on the resonance term (Problem 2):

$$\alpha_q = e^2/m\,(\omega_q^2 - \omega^2) \qquad \ldots [9.8]$$

The polarisability of the atom in this case is determined by the formula:

$$\alpha = (e^2/m) \sum (\omega_q^2 - \omega^2)^{-1} \qquad \ldots [9.9]$$

which reduces to [*7*] when $\omega \to 0$.

Polarisability of Molecules and Crystals

In homopolar molecules and crystals, the mechanism of polarisation is analogous to the electron mechanism for the polarisability of atoms. As a first approximation, the polarisability of a non-polar substance is equal to the sum of the atomic polarisabilities.

Heteropolar molecules which contain ions have dipole moments $\boldsymbol{p}_i$, of their own, which can be attributed to the individual bonds. The self dipole moment of the molecule is equal to the vector sum of the moments which make it up.

$$\boldsymbol{p}_0 = \sum \boldsymbol{p}_i.$$

If we bring a heteropolar molecule or a crystal which contains such molecules, into an electric field, three mechanisms of polarisability may take place:

1. Electron polarisability of atoms and ions which make up the molecule or crystal;
2. Ionic polarisability, which is connected with the movement of oppositely charged ions in the field;
3. Orientation polarisability, which is connected with rotation of the molecule in the electric field.

The polarisability may thus in general be expressed as a sum of three terms.

$$\alpha = \alpha_{\text{el}}+\alpha_{\text{ion}}+\alpha_{\text{or}} \qquad \ldots [9.10]$$

Different substances may have different mechanisms of polarisation. Thus in ionic coordination crystals only the first two mechanisms occur and in non-polar substances, only the electron mechanism.

Polarisation of an Ionic Mixture

Consider a diatomic heteropolar molecule A^+B^- with ionic charges $\pm e$ and with an intrinsic dipole moment $p_0 = er_0$, where r_0 is the equilibrium distance between the ions. Application of an electric field parallel to the axis of the molecle increases the distance between the ions and leads to an increase in the dipole moment. We may determine the ionic polarisability of the molecule by means of formula [6] if we introduce the reduced mass of the molecule

$$\alpha_{\text{ion}} = \frac{e^2}{k_i} = \frac{e^2}{\omega^2}\left(\frac{1}{M_1}+\frac{1}{M_2}\right), \qquad \ldots [9.11]$$

where M_1 and M_2 are the masses of the ions, k_i is the coefficient of quasi-elastic binding of the ions in the molecule, and ω is the frequency of free oscillation of the molecule. Using a potential of the form [7.6] with $m = 1$ we obtain

$$k_i = -\frac{dF}{dr} = \frac{d^2U}{dr^2} = \frac{a}{r^3}\left[\left(\frac{r_0}{r}\right)^{n-1}(n+1)-2\right].$$

In the neighbourhood of the equilibrium position for $a = e^2$ we obtain

$$k_i = (e^2/r_0^3)(n-1)$$

and putting this into the expression for the polarisability we obtain

$$\alpha_{\text{ion}} = r_0^3/(n-1) \qquad \ldots [9.12]$$

The polarisability of an ionic mixture is thus proportional to the cube of the interatomic spacing and falls as the power of the repulsive potential increases.

When an ionic crystal is brought into an electric field a similar polarisation of the ionic displacement takes place since the sub-lattices of cations move relative to the sub-lattices of anions. For a binary crystal, for example a crystal of the NaCl type, the polarisability may be determined by means of formula [*11*] if we take ω as the frequency of the residual ray (reststrahl) which corresponds to the frequency of free vibrations of the ions in the crystal. The polarisability per unit volume of the crystal is obtained by multiplying [*11*] by the number of ionic pairs contained in 1 cm^3. Now in a crystal of NaCl type, to every two ions there belongs a volume $2r^3$, where r is the distance between nearest neighbour ions. Hence the ionic polarisability of unit volume of such a crystal will be

$$n_0\alpha_{\text{ion}} = \frac{e^2}{2\omega^2 r^3}\left(\frac{1}{M_1}+\frac{1}{M_2}\right). \qquad \ldots [9.13]$$

Orientation Polarisation

The potential energy of a dipole in an electric field

$$U = -p_0 E \cos\alpha$$

is reduced when the dipole is oriented parallel to the field. The polarisation of a substance which arising from orientation polarisability is determined by the expression

$$P_{\text{or}} = n_0 p_0 \overline{\cos\alpha}, \qquad \ldots [9.14]$$

where $\overline{\cos\alpha}$ is the mean value of the cosine of the angle between the dipole and the field. Now the orienting effect of the field is opposed by thermal motion and so the mean value of the cosine is given by the Langevin function $L(x)$ (Chapter 10 §3), that is

$$\overline{\cos\alpha} = L(x) = \cot x - 1/x, \qquad \ldots [9.15]$$

where

$$x = p_0 E/kT.$$

To estimate x note that the energy of a dipole with moment $p_0 = 10^{-18}$ c.g.s. units, in a field of strength $E = 18{,}000$ V/cm $= 60$ c.g.s. units/cm, will be $p_0 E = 6\times 10^{-17}$ ergs. The mean thermal energy of a molecule at room temperature $\frac{3}{2}kT \sim 6\times 10^{-14}$ ergs, so $x \sim 10^{-3}$. When $x<1$ a good approximation to the Langevin function is the linear term

$$\cos\theta = L(x) \approx \tfrac{1}{3}x,$$

and putting this value into [*14*] we obtain the orientation polarisation

$$P_{\text{or}} = n_0 p_0^2 E/3kT \qquad \ldots [9.16]$$

and corresponding to it the orientation polarisability

$$\alpha_{\text{or}} = p_0^2/3kT. \qquad \ldots [9.17]$$

The polarisation of a substance arising from the orientation of the molecules, as distinct from the electron polarisation, depends strongly on temperature and increases as the temperature falls.

Formula [*16*] was obtained on the assumption of free rotation of the molecule and this is possible for gases and liquids. In crystals, because of the large molecular interaction, the possibility of rotation is limited, especially at low temperatures. Fig. 120 shows the temperature dependence of the dielectric constant of solid hydrogen sulphide. As the temperature falls, ε rises steeply and so consequently does the polarisability which depends on ε. This rise is connected with the influence of the orientation mechanism of polarisability since the molecules of the hydrogen sulphide have a large intrinsic moment. At 105° K there is a sharp fall in ε to a value which corresponds to the electron polarisability. This fall is a result of 'freezing' of the orientation of the dipoles in the crystal. The small discontinuities in α at higher temperatures are connected with structural changes in the crystal.

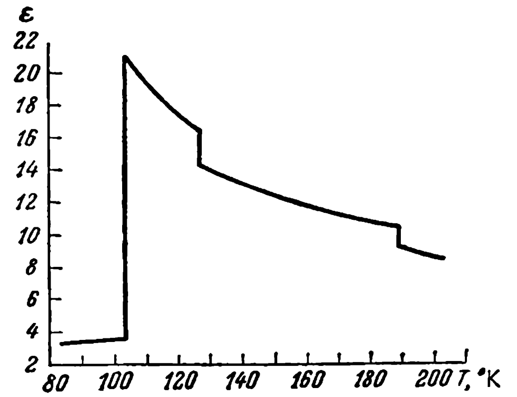


FIG. 120. Temperature dependence of the dielectric constant of solid hydrogen sulphide at a frequency of 5 kc/s.

Making the free rotations more difficult and leaving only the possibility of discrete oriented dipoles, does not change the form of the temperature dependence of the polarisability. If we assume that a dipole can have only two orientations, (*a*) parallel and (*b*) antiparallel to the field, the probability of finding it in these positions is determined by the Boltzmann expression

$$n_a/n_b = \exp(2p_0E/kT) = \exp x.$$

From which

$$\frac{\Delta n}{n} = \frac{n_a - n_b}{n_a + n_b} = \frac{e^x - 1}{e^x + 1} \approx \frac{1}{2}x = \frac{p_0E}{kT} \quad (\text{for } x \ll 1).$$

The polarisation which is connected with preferential orientation of a dipole parallel to the field is

$$P_{or} = n_0p_0^2E/kT,$$

which apart from a numerical factor agrees with [*16*].

The Internal Field

On the surface of a polarised dielectric there are surface charges, with density σ, equal to the normal component of the polarisation vector ***P***

$$\sigma = P_n.$$

The sign of the charge is negative at points where the polarisation vector enters the dielectric and positive at points where it leaves. Polarisation strengthens the field within the dielectric and helps to increase the moments of the elementary dipoles, but charges on the external surface of the dielectric reduce the field within the dielectric and have a depolarising effect. We have to calculate the internal field acting on the particles of a dielectric if we are to understand its dielectric properties and its behaviour in an electric field.

The Depolarising Field

Consider a dielectric in the form of a parallel-sided plate in a uniform electric field E, perpendicular to the surface of the plate (Fig. 121*a*). Under the influence of a field the dielectric becomes polarised. The charges of the elementary dipoles in the dielectric mutually compensate each other except for the charges on the outer surface.

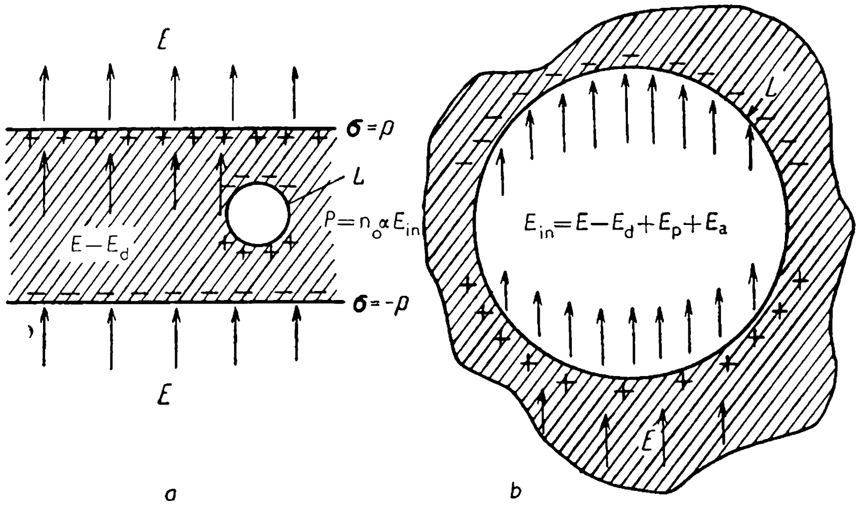


FIG. 121. Polarisation of a dielectric in a uniform electric field: (*a*) electric field within the dielectric (E the external field, E_d the depolarising field of the surface charges on the external surfaces of the dielectric); (*b*) the field within a cavity in the dielectric (E_p the polarising field of the charges on the surfaces of the cavity, E_a the field of the dipoles included inside the cavity.

These surface charges assumed to be distributed on parallel planes give an electric field of strength $2\pi\sigma = 2\pi P$. Now the field of two oppositely charged planes cancel each other outside the dielectric, but inside the dielectric they work together and give a depolarising field,

$$E_d = -4\pi P,$$

which acts against the polarising field.

The depolarising field depends on the shape of the specimen and the direction of polarisation. Uniform polarisation is possible for specimens which have the shape of an ellipsoid including also its limiting forms. If E is parallel to an axis of symmetry of the ellipsoid then the depolarising field is proportional to the polarisation of the dielectric

$$E_d = -\gamma P, \qquad \ldots [9.18]$$

where γ is the depolarising factor (for magnets this is called the demagnetising factor). For plane parallel plates $\gamma = 4\pi$, while for an ellipsoid the sum of the depolarising factors for the principal axes $\gamma_a+\gamma_b+\gamma_c = 4\pi$. Consequently for a sphere $\gamma = 4\pi/3$ (Problem 3). The depolarising factor depends on the ratio of the lengths of the axes of the specimen $l_{\parallel}/l_{\perp}$, where $l_{\parallel}$ is parallel to the field and $l_{\perp}$ is perpendicular to it and falls as this ratio increases (Table 49).

Table 49

The Depolarising Factor for Specimens of Various Shapes

Shape of the specimen	Direction of the field	$l_{\parallel}/l_{\perp}$	γ
Thin plate	$\perp$ to the plane of the plate	0	4π
Long circular cylinder	$\perp$ to the axis of the cylinder	0	2π
Sphere	Any direction	1	$4\pi/3$
Long circular cylinder	$\parallel$ to the axis of the cylinder	∞	0
Thin plate	$\parallel$ to the plane of the plate	∞	0

The Field Acting on the Particles of the Dielectric

The field which acts on the particles of the dielectric is determined in the first place by the influence of nearest neighbours and as such depends on the molecular structure of the dielectric. Knowing the atomic and molecular structure and the magnitude of the elementary dipoles we can find the resultant field at any point of the dielectric by superposing the fields of all the elementary dipoles. The general method of solution of this problem will be given later, but for the moment we shall consider a simplified method due to Lorentz: We describe a small closed surface round the point of the dielectric we are considering (Fig. 121*b*). The dimensions of this enclosure are taken such that the dielectric outside the enclosure can, with sufficient accuracy, be thought of as a continuous polarised medium while the part of the dielectric within the enclosure requires to have its atomic structure taken into account. The field inside the enclosure E_{in} can be expressed as the sum of the following fields:

$$E_{\text{in}} = E+E_d+E_p+E_a, \qquad \ldots [9.19]$$

where E is the external field, E_d the depolarising field on the external surface of the dielectric, $E_p = \gamma P$, the polarising field of the charges on the internal surface of the enclosure (γ is the depolarising factor for the shape of the enclosure), and E_a is the field of the dipoles which lie within the enclosure. If we consider a dielectric in the form of an infinitely long cylinder and the enclosure in the form of a small sphere (the Lorentz sphere) we have

$$E_{\text{in}} = E+\frac{4\pi}{3}P+E_a. \qquad \ldots [9.20]$$

For gases, isotropic liquids and amorphous dielectrics, and also for some cubic crystals, $E_a = 0$ as will be shown later (see Problems 4 and 5).

To obtain the internal field in the general case we have to find the field E_a for a system of dipoles.

The Field of a System of Point Dipoles

The dipoles in a system can be considered as points if the length of each dipole is small in comparison with the distances between them. A dipole system is determined by two families of vectors $\boldsymbol{r}_j$, and $\boldsymbol{p}_j$ ($j = 1, 2 \ldots N$), where N is the number of dipoles. The first family comprises all the radius vectors $\boldsymbol{r}_j$ which determine the position of the centres of the dipoles relative to the coordinate system XYZ (Fig. 122) which we take as right-angled. The second family comprises all the vectors $\boldsymbol{p}_j$ which describe the magnitude and direction of the dipole moments. Consider an arbitrary point K, the position of which is determined by the radius vector r_k. The field at this point is equal to the vector sum of the fields of all the dipoles (cf. Problem 4).

$$E(r_k) = \sum_{j=1}^{N} E_j(r_{kj}) = \sum_{j=1}^{N} \frac{3r_{kj}(p_j \cdot r_{kj}) - p_j r_{kj}^2}{r_{kj}^5}; \qquad \ldots [9.21]$$

$$r_{kj} = r_j - r_k.$$

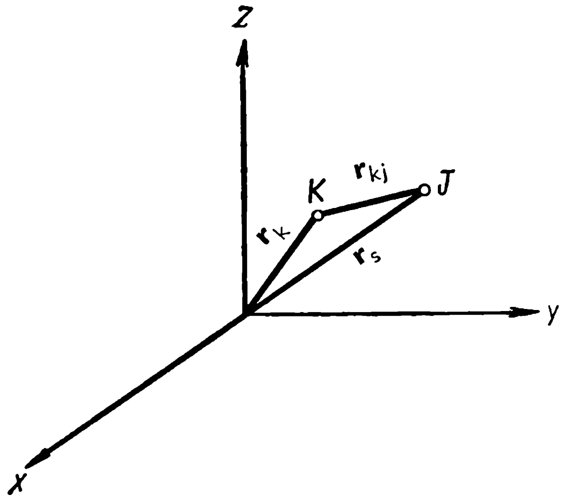


Fig. 122. Calculation of the field of a system of dipoles.

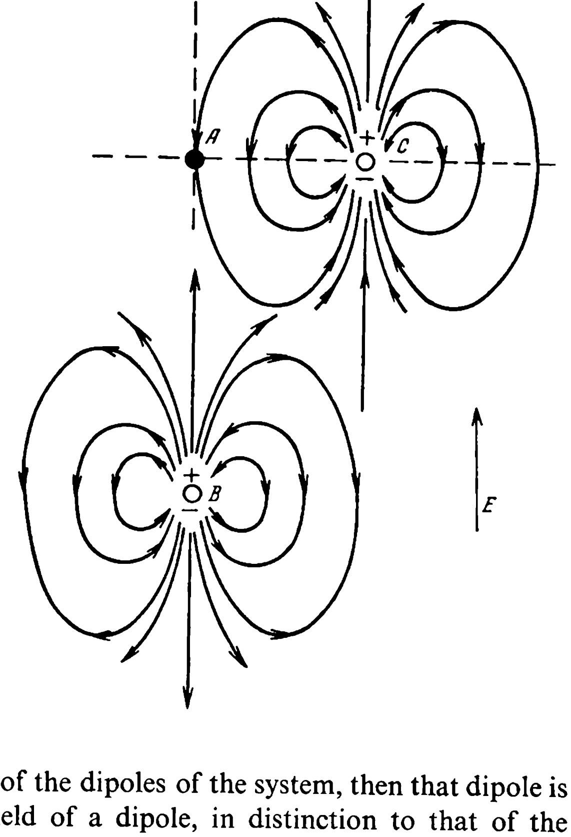
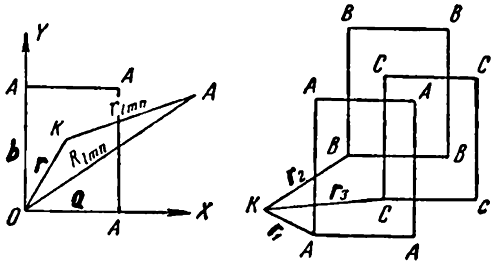


Fig. 123. Symmetry of the dipole field and superposition of the fields of two dipoles.

If the point K coincides with one of the dipoles of the system, then that dipole is omitted from the summation. The field of a dipole, in distinction to that of the centrally symmetric field of a point charge, has axial symmetry (Fig. 123). Thus sum [*21*], in general depends not only on the character of the microstructure but also on the external configuration of the system. If this is not specifically stipulated then we may take it that the system has the form of a sphere.

Let us now turn our attention to the crystalline dipole structure. In the simplest case, the unit cell contains one dipole. As a result of translational symmetry determined by the system of parallel vector translations $\boldsymbol{R}_{lmn} = l\boldsymbol{a}+m\boldsymbol{b}+n\boldsymbol{c}$ (Fig. 124*a*), all the dipoles are parallel. The direction of the dipole moments is determined by the unit vector $\boldsymbol{l}_1$. For a primitive crystal we can express sum [*21*] in the form



a *b*

FIG. 124. Calculation of the field in a crystal: (*a*) simple structure, (*b*) complex structure.

$$E(\boldsymbol{r}_k \,.\, \boldsymbol{l}_1) = pS(\boldsymbol{r}_k \,.\, \boldsymbol{l}_1),$$

where

$$S(\boldsymbol{r}_k, \boldsymbol{l}_1) = \sum_l \sum_m \sum_n \frac{3(\boldsymbol{R}_{lmn}-\boldsymbol{r}_k)(\boldsymbol{l} \cdot \boldsymbol{R}_{lmn}-\boldsymbol{r}_k)-\boldsymbol{l}_1(\boldsymbol{R}_{lmn}-\boldsymbol{r}_k)^2}{|\boldsymbol{R}_{lmn}-\boldsymbol{r}_k|^5} \quad \ldots [9.22]$$

The quantity $S(\boldsymbol{r}_k \,.\, \boldsymbol{l}_1)$ depends only on the geometry of the system and is called the structural sum of the dipole lattice. The components of S along the coordinate axis S_x, S_y, S_z are called *structural coefficients*. Resolving the polarisation vector $p\boldsymbol{l}_1$ into components p_x, p_y and p_z we obtain the structural coefficients corresponding to the directions of polarisation along the coordinate axes. Knowing these structural coefficients we can find the field for any direction of polarisation. For example, for a cubic crystal

$$S_x^{110}(xyz) = \frac{1}{\sqrt{2}}\{S_x^{100}(xyz)+S_x^{010}(xyz)\}.$$

Table 50 gives expressions for the structural coefficients for a cubic lattice for the three directions of the polarisation vector which coincide with the directions of the axes of symmetry of the cube. Values of the structural coefficients are given for the principal points of the unit cell. For the vertices and centre of the cell the structural coefficients are zero; for the face centres they are determined by the parameter Q and for the centres of the edges by the parameter P. In a cubic lattice these parameters are inversely proportional to the volume of the unit cell. The numerical values of the parameters determined by various methods are in good agreement. By the Ewald method McKeehan obtained

$$P = -15{\cdot}04102/a_0^3; \quad Q = 4{\cdot}33387/a_0^3, \quad \ldots [9.23]$$

where a_0 is the length of the edge of the unit cube.

For a crystal with several particles in the unit cell, the subgroup of parallel translations is the same for all particles. The structure of such a crystal can be thought of as a superposition of sub-lattices A, B, C . . . each containing particles of one kind and displaced relative to each other in parallel directions (Fig. 124*b*). Here the field at the point K is equal to the sum of the fields of the sub-lattices

$$E(r_k) = \sum_i p_i S(r_{ik}), \qquad \ldots [9.24]$$

where the structural sums are the same for all the sub-lattices. Knowing the structural coefficients we can find the internal field at different points of the crystal.

Introducing into [*24*] the polarisability of the dipoles we obtain

$$E_k = \sum_i \alpha_i S(r_{ik}) E_i, \qquad \ldots [9.25]$$

where E_i is the internal field acting on the ion i. By shifting the point K into coincidence with the various ions we obtain a system of simultaneous linear equations for the internal fields E_k, E_i acting on the ions. In the absence of an external field

Table 50

Structural Coefficients for a Simple Cubic Dipole Lattice

x, y, z	p	$S_p\,(x, y, z;\ l_1)$				
		[100]	[010]	[001]	[110]	[111]
$0\ 0\ 0$	x	0	0	0	0	0
$\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}$	y	0	0	0	0	0
	z	0	0	0	0	0
$\frac{1}{2}\ \frac{1}{2}\ 0$	x	Q	0	0	$Q/\sqrt{2}$	$Q/\sqrt{3}$
	y	0	Q	0	$Q/\sqrt{2}$	$Q/\sqrt{3}$
	z	0	0	$-2Q$	0	$-2Q/\sqrt{3}$
$\frac{1}{2}\ 0\ \frac{1}{2}$	x	Q	0	0	$Q/\sqrt{2}$	$Q/\sqrt{3}$
	y	0	$-2Q$	0	$-2Q/\sqrt{2}$	$-2Q/\sqrt{3}$
	z	0	0	Q	0	$Q/\sqrt{3}$
$0\ \frac{1}{2}\ \frac{1}{2}$	x	$-2Q$	0	0	$-2Q/\sqrt{2}$	$-2Q/\sqrt{3}$
	y	0	Q	Q	$Q/\sqrt{2}$	$Q/\sqrt{3}$
	z	0	0	Q	0	$Q/\sqrt{3}$
$0\ 0\ \frac{1}{2}$	x	P	0	0	$P/\sqrt{2}$	$P/\sqrt{3}$
	y	0	P	0	$P/\sqrt{2}$	$P/\sqrt{3}$
	z	0	0	$-2P$	0	$-2P/\sqrt{3}$
$0\ \frac{1}{2}\ 0$	x	P	0	0	$P/\sqrt{2}$	$P/\sqrt{3}$
	y	0	$-2P$	0	$-2P/\sqrt{2}$	$-2P/\sqrt{3}$
	z	0	0	P	0	$P/\sqrt{3}$
$\frac{1}{2}\ 0\ 0$	x	$-2P$	0	0	$-2P/\sqrt{2}$	$-2P/\sqrt{3}$
	y	0	P	0	$P/\sqrt{2}$	$P/\sqrt{3}$
	z	0	0	P	0	$P/\sqrt{3}$

equations [*25*] describe a state of spontaneous polarisation. The condition for solubility of equations [*25*] is that the following determinant should be zero.

$$\begin{vmatrix} S_{11}\alpha_1-1, & S_2, \alpha_2, \ldots, & S_{n1}\alpha_n \\ S_{12}\alpha_1, & S_{22}\alpha_2-1, \ldots, & S_{n2}\alpha_n \\ \mathrm{S}_{13} \;\cdot & \cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot\;\cdot \\ S_{1n}\alpha_1, & S_{2n}\alpha_2, \ldots, & S_{nn}\alpha_n-1 \end{vmatrix} = 0. \qquad \ldots [9.26]$$

For a given composition and crystal structure the polarisabilities α_i and the structural coefficients S_{ik} can be taken as known. Relations [*26*] then describe the connection between the properties of the particles (α_i) and the structure (S_{ik}) when there is spontaneous polarisation.

The Energy of a System of Point Dipoles

The energy of a dipole in an electric field is

$$u = -(\boldsymbol{p}\boldsymbol{E}).$$

The total potential energy of a system of dipoles is

$$U = -\tfrac{1}{2}\sum_j \sum_k u_{jk} = -\tfrac{1}{2}\sum_j \sum_k (\boldsymbol{p}_j \cdot \boldsymbol{E}_k). \qquad \ldots [9.27]$$

The sum [*27*] refers to unit volume of the system. The energy of a dipolar crystal which contains n_0 unit cells in unit volume will be

$$U = -\tfrac{1}{2} n_0 \sum_j \sum_k (\boldsymbol{p}_j \boldsymbol{E}_k),$$

where the summation over j is carried out over all the particles within the unit cell and the summation over k is carried out over all the sub-lattices of the crystal. For a primitive cubic crystal polarised along the Z axis (cf. Problem 4) the potential energy can be expressed in the form

$$U = -\tfrac{1}{2} n_0^2 p^2 \sum_{l,m,n} \frac{2l^2-m^2-n^2}{(l^2+m^2+n^2)^{5/2}}, \qquad \ldots [9.28]$$

where l, m, and n are integer coordinates.

The energy of a dipole system, like its field strength depends on the external shape of the system and the direction of the polarisation vector. For an ellipsoidal specimen polarised along an axis of symmetry of the ellipsoid

$$U = -\tfrac{1}{2}\sum p_j(\tfrac{4}{3}\pi-\gamma)p = -\tfrac{1}{2}\left(\frac{4\pi}{3}-\gamma\right) n_0^2 p^2, \qquad \ldots [9.29]$$

where γ is the depolarising factor. The energy of a uniformly polarised system does not depend on the type of lattice but does depend on the shape of the specimen. For specimens of spherical shape $U = 0$. As the axial ratio of the ellipsoid c/a increases, the potential energy of the dipole system falls and reaches a minimum:

$$U_{\min} = -\frac{2\pi}{3} n_0^2 p^2 = -2{\cdot}1\, n_0^2 p^2 \qquad \ldots [9.30]$$

for specimens having the shape of an infinitely long cylinder. For long specimens the energy of the polarised state is less than the energy of a random distribution of dipoles

($U = 0$). If the dipoles are free to rotate we may expect that spontaneous polarisation will arise if there are no other dipole configurations which have an energy less than [*30*].

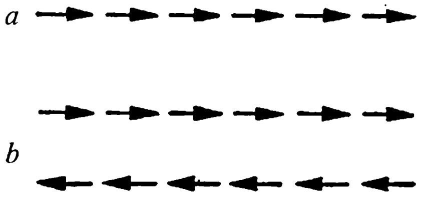


FIG. 125. (*a*) Uniform chain of dipoles, (*b*) antiparallelism of polar chains.

It is not difficult to find the dipole structure which has the lowest potential energy in one- and two-dimensional lattices: in the one-dimensional case the minimum energy is that of a polarised state such that the dipoles lie parallel along the axes of the dipole chain (Fig. 125*a*); for two parallel chains and a plane network, the minimum energy is that of the antipolar structure (Fig. 125*b*) with antiparallel arrangement of the linear polar chains.

The case of a three-dimensional crystal is much more difficult. In the atomic structure of such a crystal there can exist a countless number of dipole structures distinguished from each other by the relative orientations of the dipoles. Consider antipolar structures in which the dipoles are situated relative to each other in two positions only: parallel (A) and antiparallel (B). If we take the number of dipoles in both positions to be the same then the resultant moment will be zero. This means that in a given dipole structure it is necessary to investigate superlattices of the AB type each of which is distinguished by the orientation of its dipoles.

In the simple cubic lattice the simplest superlattice of interest to us is that in which the antiparallel polar chains are situated according to a chess board pattern (Fig. 126*a*). The unit cell of this superlattice is tetragonal. The potential energy

$$U(\Gamma_c) = -2{\cdot}7\, n_0^2 p^2 \qquad \ldots [9.31]$$

is less than that given by equation [*30*] so the antipolar configuration is more stable than the polar structure. Thus spontaneous polarisation is impossible in a system of freely rotating dipoles situated at the corners of a simple cubic lattice, even if the system has the shape of an infinitely long cylinder.

In a b.c.c. lattice the simplest dipole superlattice is shown in Fig. 126*b*. Its energy $U = 0$. A somewhat more complex superlattice (Fig. 126*c*) has energy $U = -1{\cdot}35$, but there is a still smaller energy for a superlattice which differs from it in the directions of the dipoles: in the former case the dipoles are directed along the edges of the cube while in the latter they lie along the body diagonals, that is along the shortest lattice spacings. The change in direction of the dipoles gives an essential change in the symmetry of the superlattice and in the shape and volume of the unit cell which, from being tetragonal and containing four dipoles in the former case, is now rhombic with six dipoles. The energy of this superlattice type (Fig. 126*d*)

$$U(\Gamma_c') = -1{\cdot}75\, n_0^2 p^2 \qquad \ldots [9.32]$$

can be shown to be a minimum.

Among f.c.c. structures the simplest point superlattice AB is tetragonal and of the AuCu type. Figs. 126*d*, *e*, and *f* show three dipole superlattices which are formed on the basis of this point superlattice and differ from each other in the directions of orientation of the dipoles: [001], [100] and [110]. Changes in the orientations lower the symmetry of the dipole structure from tetragonal in the first case to rhombic in the other two cases. The energy of the first structure, in which nearest chains are parallel, is $U = 2{\cdot}2$ which is large and exceeds the energy of random structures. In the structure of Fig. 126*e* nearest chains are antiparallel and this change of orientation of the dipoles markedly lowers the energy to $U = -1{\cdot}1$. However though the positions of neighbouring chains are favourable from the point of view of a gain in energy of interaction between them, the dipoles are not directed along the shortest lattice spacing. Fig. 126*f* shows the simplest antipolar superlattice with directions of the dipole axes along the base diagonals. Here the interactions of the dipoles in the chains is more favourable but the energy of the structure comes out to be $U = -1{\cdot}1$, the same as in the preceding case so this distribution of neighbouring chains does not satisfy optimum conditions. Fig. 126*h* shows a superlattice which does satisfy the desired optimum distribution, both for dipoles in chains and for the chains relative to each other. The unit cell of this lattice is rhombic and contains eight dipoles. Its energy

$$U(\Gamma_c'') = -1{\cdot}8\, n_0^2 p^2 \qquad \ldots [9.33]$$

is a minimum.

The energy of the most favourable antipolar structures in the b.c.c. [*32*] and f.c.c. [*33*] lattices exceed the energy of the polar structure [*30*]. Consequently in these lattices, when the shape of the specimen is favourable, it is in principle possible to obtain spontaneous polarisation. In dipole systems with the f.c.c. structure, spontaneous polarisation can arise in ellipsoidal specimens with axial ratio exceeding 6·7.

A warning is necessary about using the calculations given above for magnetic materials: in ferromagnetic metals and alloys the main factor in the development of spontaneous magnetisation is the exchange interaction which dies away rapidly as the distance between the elementary carriers of magnetism is increased, and indeed the model of a ferromagnetic metal is a system of polar particles in which only nearest neighbours interact (the Ising Model) and are oriented relative to each other in only two positions ↑↑ or ↓↑. In magnetic dielectrics (ferrites) the magnetically active ions are separated by inactive ions and the exchange interaction is less important. In this case, we may use the dipole model to obtain the energy of dipolar interaction.

The Dielectric Constant and the Polarisability

One of the principal characteristics of a dielectric is the dielectric constant ε. In isotropic media ε is a scalar and relation [*2*] takes a simple form; the vector of induction is parallel to the field strength. In anisotropic media the dielectric constant depends on the direction of E. In the general case the components of the induction vector are related to the components of the field strength by relations of the form

$$\left.\begin{aligned} D_x &= \varepsilon_{11}E_x+\varepsilon_{12}E_y+\varepsilon_{13}E_z \\ D_y &= \varepsilon_{21}E_x+\varepsilon_{22}E_y+\varepsilon_{23}E_z \\ D_z &= \varepsilon_{31}E_x+\varepsilon_{32}E_y+\varepsilon_{33}E_z \end{aligned}\right\}. \qquad \ldots [9.34]$$

These show that the dielectric constant is a tensor of the second rank. By applying the law of conservation of energy we can show that $\varepsilon_{ik} = \varepsilon_{ki}$ (cf. [*15.15*]), so the tensor (ε_{ik}) is symmetrical and is represented by an ellipsoid. In a system of coordinates which coincide with the directions of the principal axes of the ellipsoid, the non-diagonal components reduce to zero and the tensor is determined by the three diagonal components $\varepsilon_{11} = \varepsilon_1$; $\varepsilon_{22} = \varepsilon_2$; $\varepsilon_{33} = \varepsilon_3$. For example in the rhombic crystal $PbSO_4$, $\varepsilon_1 = 27{\cdot}5$; $\varepsilon_2 = 54{\cdot}6$; $\varepsilon_3 = 27{\cdot}3$ and the corresponding ellipsoid is stretched out along the *b* axis which is the shortest axis of the crystal. In uniaxial crystals the triaxial ellipsoid reduces to an ellipsoid of revolution. Thus in the hexagonal crystal of α-quartz $\varepsilon_1 = \varepsilon_2 = 4{\cdot}49$; $\varepsilon_3 = 4{\cdot}55$. In cubic crystals and isotropic media ε does not depend on direction and is represented by a sphere. To simplify the further analysis we shall as a rule assume that we are dealing with an isotropic medium.

By definition, the polarisation is

$$P = \sum_i n_i \alpha_i E_i. \qquad \ldots [9.35]$$

For a dielectric which consists of one kind of molecule

$$P = n_0 \alpha E_{in}, \qquad \ldots [9.36]$$

where n_0 is the number of molecules in unit volume. Assuming that the internal field is determined by [20] where $E_a = 0$ we obtain from [20], [36] and [3]

$$P/E = n_0\alpha/(1 - \tfrac{3}{4}\pi n_0 \alpha) = (\varepsilon - 1)/4\pi.$$

Hence we obtain the relation between the dielectric constant and the polarisability of the molecules of the dielectric:

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \frac{4\pi}{3} n_0 \alpha. \qquad \ldots [9.37]$$

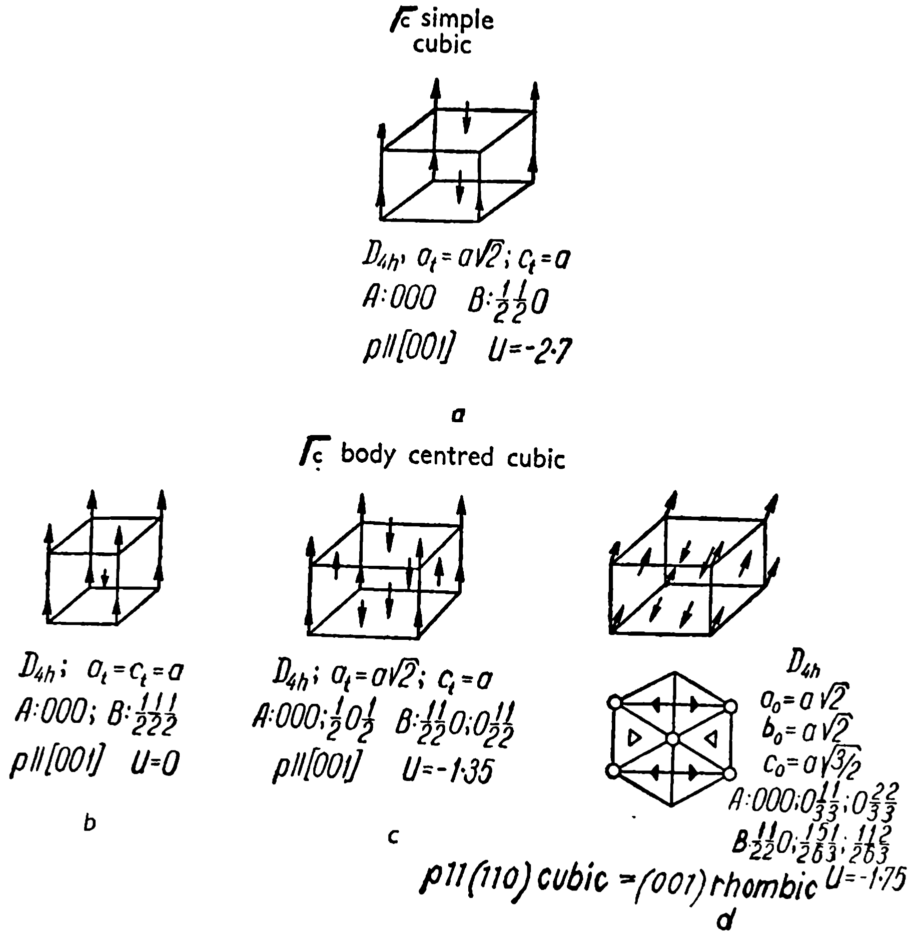


FIG. 126. The simplest cubic

Multiplying both sides by the molecular volume

$$V = \frac{M}{\rho}$$

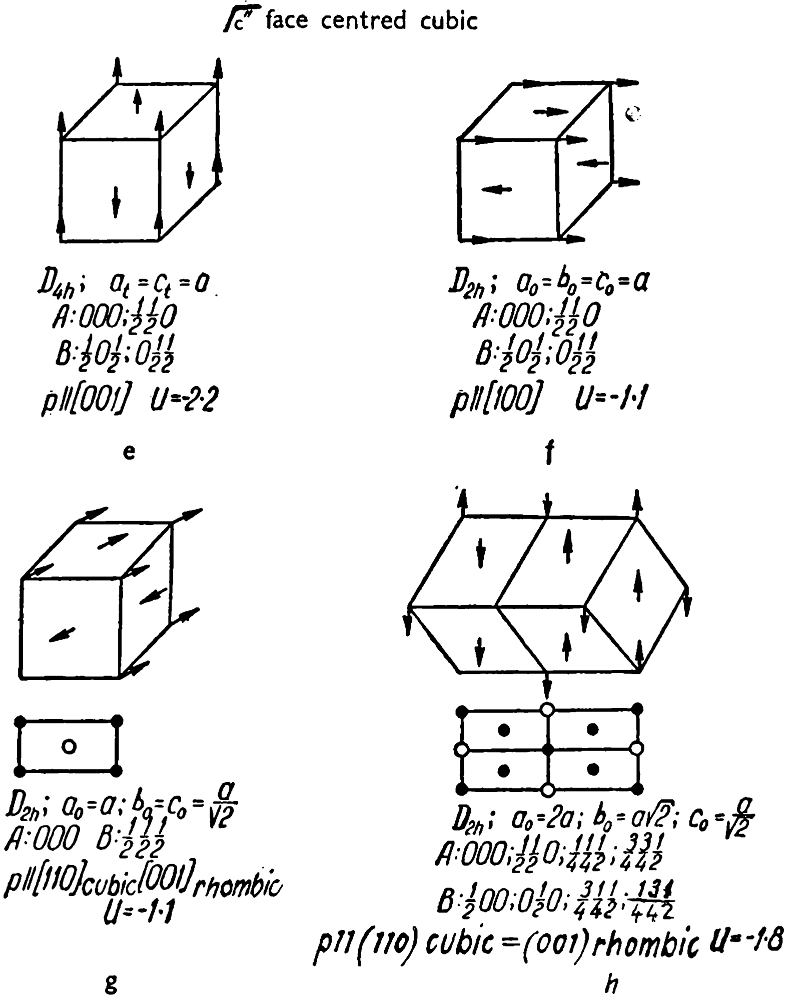
(where M is the gm-molecular weight and ρ is the density of the dielectric) and remembering that $Vn_0 = N$, where N is Avogadro's number, we obtain

$$\frac{M}{\rho}\frac{\varepsilon-1}{\varepsilon+2} = P_M, \qquad \ldots [9.38]$$

where

$$P_M = \frac{4\pi}{3} N\alpha \qquad \ldots [9.39]$$

the polarisability of a gram-molecule of the substance. This is called the molecular polarisation. Equation [*38*] which is well-known under the name of the Clausius-Mosotti formula, expresses the relation between the dielectric constant, the molecular weight, the density and the molecular polarisation. Measurements of the dielectric



anti-polar dipole super-lattices.

constant thus enable us to determine the molecular polarisation P_M and the polarisability α of the individual molecules.

The electromagnetic theory of light establishes the relation between the dielectric constant and the refractive index of a medium

$$n^2 = \varepsilon. \qquad \ldots [9.40]$$

Putting this relation into [*38*] we obtain the Lorenz-Lorentz law

$$\frac{M}{\rho}\frac{n^2-1}{n^2+1} = R, \qquad \ldots [9.41]$$

which determines the relation between the refractive index of a medium and the molecular refraction

$$R = \frac{4\pi}{3} N\alpha. \qquad \ldots [9.42]$$

The polarisability of a molecule α as determined from dielectric measurements can be quite different from α as determined by optical measurements since the polarisability depends on the frequency of oscillation (see [*9*]).

Fig. 127 shows the dependence of the molecular polarisation of a polar dielectric on the frequency of the electric field. Over a considerable range of frequencies, from the static field through radio frequencies right to the infra-red, the molecular polarisation does not depend on frequency and is large. As the frequency is further increased into the ultra-violet region, the polarisation falls steeply. We can distinguish three regions on the curve, establishing the existence of three different mechanisms of polarisation. As a result of the difference of mass of the electrons and ions only electron polarisation is possible in the high frequency region corresponding to ultra-violet radiation. The frequencies of vibrations of atoms lie within the infra-red region and from there down, ionic polarisation can take place. Orientation of molecules in the periodic field becomes possible at still lower frequencies, corresponding to centimetre and metre waves.

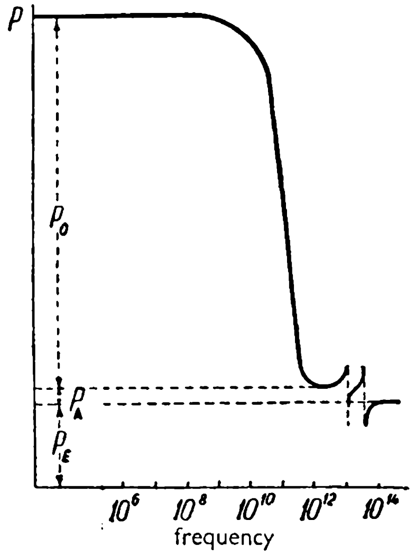


FIG. 127. Dependence of the molecular polarisation of a polar dielectric on the frequency of the field.

Polarisability and Refraction of Some Substances

Table 51 gives the relation between the square of the refractive index and the dielectric constant for various crystalline substances. The Table shows that these substances can be divided into three groups.

Table 51

Values of n^2 and ε for Some Crystals

Crystal	n^2	ε	Mechanism of Polarisation
Rhombic sulphur ‖ principal optic axis	5·0	5·06	Electronic
Rhombic sulphur ⊥ principal optic axis	3·8–4·15	4·0	
Naphthalene	2·50	2·50	
NaCl	2·35	5·6	Electronic and ionic
KCl	2·22	4·9	
α–quartz	2·35	4·5	
Fused quartz		3·5	
Ice	1·71	94	Electronic ionic and orientation
Water	1·77	81 $\lambda = \infty$ 3·3 λ = 3600 cm 2·8 λ = 1200 cm	

To the first group belong molecular crystals of the sulphur and naphthalene type with homopolar bonds in the molecules. For these crystals the relation $n^2 = \varepsilon$ is in good agreement with observation. To the second group belong ionic coordination crystals for which ε is somewhat greater than n^2; this is accounted for by polarisation of ionic movement. In the third group are polar crystals, for example ice. Among minerals, ice has one of the smallest refractive indices but nevertheless large compared with that of water. As the temperature is lowered, ε for ice falls (Fig. 128) and tends to the value for non-polar ionic crystals. The temperature region in which the rapid fall of ε occurs, depends on the frequency of the field and moves towards the region of low temperature as the frequency is lowered.

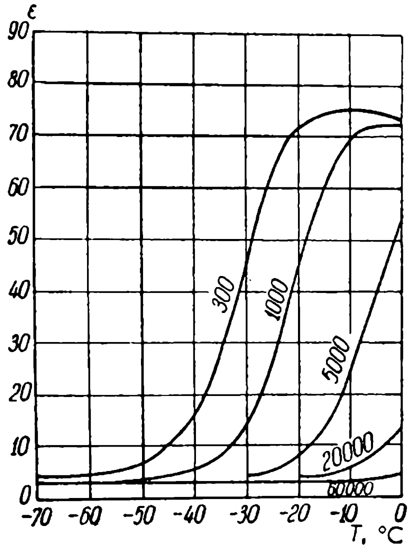


FIG. 128. Dependence of the dielectric constant of ice on the temperature and frequency of the field (in cycles/sec).

In the region of large values of ε, an increase in the frequency at constant temperature leads to a reduction in the dielectric constants of water and ice. Our explanation shows that orientational polarisation is the dominant mechanism in the dielectric constant of water and ice only if it is not excluded either on account of the inertia of the molecules at high frequencies or on account of an increase of the viscosity at low temperatures. Breakdown of a regular crystal structure as can be seen by the example of fused (amorphous) quartz and the melting of ice leads to a reduction of ε.

According to Debye the orientational polarisability for polar liquids and solutions is a complex quantity

$$\alpha_{or} = \alpha_0/(1+i\omega\tau), \qquad \ldots [9.43]$$

where ω is the frequency of the field and τ is the relaxation time proportional to the viscosity of the medium η. For spherical molecules of radius a

$$\tau = 4\pi\eta a_0^3/kT.$$

There is a special group of substances which have a large value of ε but in which the orientation effect is missing. To this group also belong substances with anomalously high values of ε and non linear characteristics (ferroelectrics; see below). An example of a coordination ionic crystal with a large value of the dielectric constant is rutile for which the mean value of $\varepsilon = 114$. Other modifications of titanium dioxide —brookite and anatase—also show large values of ε—78 and 48 respectively. The reduction of ε in polymorphic modifications of titanium dioxide may be compared with the corresponding reduction in their density: 4·21; 4·11; 3·87.

Atomic and Ionic Polarisability

Table 52 gives values of the polarisability of some atoms, determined from molecular refraction and molecular polarisability.

Table 52

Polarisability of Atoms in Homopolar Compounds

$\alpha\times10^{24}$ cm^3			H 0·1	He 0·20
C 0·96	N 0·96	O 0·64	F 0·4	Ne 0·39
			Cl 2·4	Ar 1·65
			Br 3·6	Kr 2·51
			I 5·8	Xe 4·1

Table 53

The Polarisabilities of Ions in Heteropolar Compounds

$\alpha\times10^{24}$ cm^3

Li^+ 0·079	Be^{+2} 0·035							B^{+3} 0·020	C^{+4} 0·012		O^{-2} 2·76	(
Na^+ 0·197	Mg^{+2} 0·114							Al^{+3} 0·067	Si^{+4} 0·039		S^{-2} 5·90	
K^+ 0·879	Ca^{+2} 0·531	Sc^{+3} 0·382	Ti^{+4} 0·272			Cu^+ 1·81 Cu^{+2} 0·67	Zn^{+2} 0·114				Se^{-2} 6·42	
Rb^+ 1·81	Sr^{+2} 1·42	Y^{+3} 1·02	Zr^{+4} 0·800			Ag^+ 1·85	Cd^{+2} 0·96				Te^{-2} 9·60	
Cs^+ 2·48	Ba^{+2} 1·69	La^{+3} 1·58					Hg^{+2} 1·99		Pb^{+2} 4·82			

As the atomic number increases in a series of homologous compounds there is an increase in the atomic polarisability. This is explained by the increase in the dimensions of the atoms and reduction in the strength of the valence electron binding as shown by the fall in ionisation potentials. Within a single period, the polarisabilities fall slightly; this is connected with an increase in the nuclear charge. The polarisabilities of positive ions (Table 53) are smaller and of negative ions larger than the polarisabilities of the corresponding atoms on account of the increase of the deformability of the electronic cloud with increase in the number of electrons at a constant nuclear charge.

3. Crystalline Electrets and Pyroelectrics. Piezoelectrics

The Symmetry Classes of Polar Dipole Lattices

Polar directions cannot exist in crystals with a centre of symmetry. Such crystals cannot therefore have a constant polarisation and are not pyro- or piezo-electrics. Table 54 shows point groups and crystalline classes without a centre of symmetry. The formation of polar dipole lattices is possible only in groups with a polar axis, that is in cyclic groups C_n and in axial groups with a vertical plane of symmetry C_{nv}. The resultant dipole moment is then directed along the polar axis. In the monoclinic class $C_{1v}-m$, which contains only a plane of symmetry, the resultant moment lies in that plane. In the triclinic class C_1 which has no symmetry elements, the resultant moment can have any direction relative to the crystallographic axes. To the point groups C_n and C_{nv} there belong ten crystal classes in which symmetry considerations make an electret state and pyroelectricity possible.

Crystalline Electrets

Some substances and their corresponding symmetry classes are given in Table 54. Whether or not crystals of the first ten classes have constant polarisation depends not only on symmetry, which is a necessary but not a sufficient requirement, but also on the properties of the particles which form the crystal. Thus in the molecular crystal of hexamethyl benzole which belongs to class C_1, the hexagonal molecules $C_6(CH_3)_6$ are non-polar so the resultant dipole moment of the crystal is zero.

If the dipole moment of the unit cell differs from zero, then the crystal as a whole will be polarised and will create a field of force around it with magnitude determined by the surface density σ_n of the dipole charges on the external surface of the crystal. It is difficult to demonstrate this electrical polarisation experimentally since the surface charges are neutralised by charged particles and ions attracted from the surrounding medium. However one method of demonstrating the polarisation consists in breaking the crystal into two parts and investigating the charges on the fracture surfaces. Thus, Foëx broke needle-shaped crystals of tourmaline and measured a surface density of charge given by $\sigma_n = 33$ c.g.s. units/cm^2. It would be better however to measure the changes of polarisation of the crystal under the influence of external isotropic factors such as temperature or hydrostatic pressure.

Table 54

Point Groups and Crystal Classes not having Centres of Symmetry

No.	Point group	Crystal class	Symmetry	Examples	Possible properties	
1	C_n	C_1-1	triclinic	$CaS_2O_3 \cdot 6H_2O$; $C_6(CH_3)_6$ – hexamethyl benzol $Rb_4Fe(CN)_6 \cdot 2H_2O$; $C_6H_4(CH_3)NH \cdot C(CH_3)$	PIEZO	crystalline electrets,
2		C_2-2	monoclinic	MoO_2; α-Pb_2O_3 $H_2C_4H_4O_6$ – tartaric acid		
3		C_3-3	hexagonal	Tl_2S; B_2O_3 $Na_2I_2O_8 \cdot 6H_2O$		
4		C_4-4	tetragonal	$PbMoO_4$—wulfenite		
5		C_6-6	hexagonal	BCl_3; PI_3; $Sr(SbO)_2(C_4H_4O_6)_2$ $Na(AlSiO_4)$ – nepheline		
6	C_{nv}	$C_{1v}-m$	monoclinic	$K_2S_4O_6$ $C_6H_4(CH_3)NHC(CH_3) \cdot CO_2C_2H_5$		
7		$C_{2v}-mm$	rhombic	V_2O_5; β-Pb_2O_3; SO_2; Sb_2O_4 $(NH_4)MgPO_4 \cdot 6H_2O$; $Zn_2H_2SiO_5$ – calamine		
8		$C_{3v}-3m$	hexagonal	P_2O_5; γ-NiS; $SiCl$; $C_{14}H_{15}O$ – tolylphenyl ketone tourmaline; $(C(NH_2)_3)(M(H_2O)_6)(EO_4)_2$ $M = Al, Cr$, $E = S, Se$		
9		$C_{4v}-4m$	tetragonal	β-U; B_6H_9; $C_4H_4O_2 \cdot NI$ $C(CH_2OH_4)$ – pentaerythritol		

pyroelectrics, and ferroelectrics

10		$C_{6v}-6m$	hexagonal	β-ZnS – wurtzite: CdI_2 4-layer type $Ca_5(PO_4)X$; $X = F, Cl, OH$ – hepatite	ELECTRICS
11	S_{2n}	$S_4-\bar{4}$	tetragonal	$Ca_2Al_2SiO_7$	
12	C_{nh}	$C_{3h}-\bar{6}$	hexagonal	Ag_2HPO_4	
13	D_n	D_2-222	rhombic	$MgSO_4 \cdot 7H_2O$ γ-$NaKC_4H_4O_6 \cdot 4H_2O$ – Rochelle salt	
14		D_3-32	hexagonal	γ-Se, Te, HgS, Fe_2P α-quartz	
15		D_4-42	tetragonal	β-Np; $CCl_3(CO_3K)CCl_3(CO_2H)$ – potassium trichloracetate	
16		D_6-62	hexagonal	$K_1MoO_{12}SiO_{40}$ β-quartz	
17	D_{nh}	$D_{3h}-\bar{6}m2$		$BaSi_3TiO_9$ – benitoite	
18	D_{nd}	$D_{2d}-\bar{4}m2$	tetragonal	B; $Hg(CN)_2$; $CuFeS_2$	
19	$T-23$		cubic	$NaClO_3$; $PbNO_3$	
20	$Td-\bar{4}3m$			α-ZnS – sphalerite	
21	$O-43$			β-Mn, Cu_2O; NH_4Cl	
22	$J-53$				

Pyroelectrics

As Table 54 shows, dipolar structures are possible for all crystallographic symmetry types except cubic. When the temperature of a crystal belonging to a non-isotropic system is changed, its dimensions change anisotropically as a result of changes of the interatomic spacings and the valence angles. This leads to a change of the dimensions and positions of the elementary dipoles. Consequently, the resultant moment of the crystal and the surface density of polarisation charges change. This phenomenon is called pyroelectricity. The crystalline classes of pyroelectrics are the same as for polar lattices.

Phenomenologically, pyroelectricity can be described by the relation

$$\Delta P = \Delta\sigma_n = k\,(T_2 - T_1), \qquad \ldots [9.44]$$

which expresses a linear relation between the change in electrical polarisation ΔP, and the change of temperature from T_1 to T_2 which produces it. The coefficient of proportionality k depends on the chemical composition and structure of the crystal. The pyroelectric effect occurs similarly when crystals are cooled in liquid air. In tourmaline an increase in the pyroelectric coefficient with temperature has been established:

$$k = 1{\cdot}13 + 0{\cdot}0104\,(t - 18^\circ\,\mathrm{C})\ \text{c.g.s. units.}$$

Other pyroelectrics are the salts of tartaric acid: potassium tartrate $K_2C_4H_4O_6 \,.\, \frac{1}{2}H_2O$, ammonium tartrate $(NH_4)_2C_4H_4O_6$ and Rochelle salt $NaKC_4H_4O_6 \,.\, 4H_2O$ between its Curie points, the sulphates and selenates of lithium ($Li_2EO_4 \,.\, H_2O$; E = S, Se, and also LiNa-sulphate and LiK-sulphate), wurtzite (β-ZnS) and so on.

When a pyroelectric is brought into an electric field the reverse effect is to be expected, that is a change in temperature of the crystal. In tourmaline this *electrothermal* effect is $1{\cdot}4 \times 10^{-5}$ °C at room temperature in a field of one electrostatic unit.

Piezoelectrics

Changes in the state of polarisation must also result from mechanical deformations which change the magnitude and direction of the elementary dipoles. Thus, deformation by extension or shear will change the polarisation of all dipolar structures which do not have a centre of symmetry so, to the 10 pyroelectric classes must be added another 11 crystal classes which have some polar symmetry axes (Table 54). The production of surface electric charges by deformation was first demonstrated on a crystal of quartz by Pierre Curie (1880). When a piezoelectric crystal is brought into an electric field the reverse effect takes place—the development of elastic stresses and deformation of the crystal (Problem 6).

The mechanical stress in a crystal is described by a tensor of second rank $\| t_{ik} \|$ (see [*16.5*]). The polarisation produced in the crystal is described by a vector $\boldsymbol{p}$, the components of which can be called P_l. In piezoelectricity, polarisation is connected with mechanical stress by a linear relation of the form

$$P_l = d_{ikl} t_{ik}. \qquad \ldots [9.45]$$

Piezoelectricity is described by a tensor of the third rank $\| d_{ikl} \|$, the components of which are called the piezoelectric moduli. The stress tensor is symmetrical $t_{ik} = t_{ki}$, so

$$d_{ikl} = d_{kil} \qquad \ldots [9.46]$$

The piezoelectric tensor is in general determined by 18 components and can be written in a simpler form:

$$\begin{array}{c|cccccc} & t_{11} & t_{22} & t_{33} & t_{23} & t_{31} & t_{12} \\ \hline P_x & d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ P_y & d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} \\ P_z & d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} \end{array} \qquad \ldots [9.47]$$

The stress tensor has a centre of symmetry. Thus in a uniformly stressed state, a centro-symmetric crystal retains its centre of symmetry. Since in groups with a centre of symmetry all directions are non-polar, the piezo effect is absent in crystals which have a centre of symmetry. This means that all components of the tensor $\| d_{ik} \|$ are zero. The presence of these or other symmetry elements in acentric groups imposes limits on the piezoelectric moduli. Thus the presence of a second order axis of symmetry leads to zero values for 10 components of the tensor (Problem 7). In point group 43 all components $d_{ik} = 0$ as a result of symmetry. Thus piezoelectricity is possible in 20 of the 21 classes which do not have centres of symmetry.

In the trigonal crystal of α-quartz there is one non-polar triad axis of symmetry (the optic axis) and three polar diad axes (the electric axes). In an orthogonal system of coordinates the (X-axis along one of the electric axes, the Z-axis along the optic axis) the piezoelectric tensor for group 32 has the form

$$\begin{array}{c|cccccc} & t_{11} & t_{22} & t_{33} & t_{23} & t_{31} & t_{12} \\ \hline P_x & d_{11} & -d_{11} & 0 & d_{14} & 0 & 0 \\ P_y & 0 & 0 & 0 & 0 & -d_{14} & -2d_{11} \\ P_z & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

and is determined by two independent moduli. The components of the vector of electrical polarisation produced by plastic stress in quartz are determined by the relations

$$P_x = d_{11}(t_{11}-t_{22})+d_{14}t_{23};$$
$$P_y = -d_{14}t_{31}-2d_{11}t_{12}; \quad P_z = 0.$$

For a given uniform deformation, the vector of polarisation lies in a plane perpendicular to the optic axis. For uniform linear deformation, the shear stresses

$$t_{23} = t_{31} = t_{12} = 0, \quad P_y = 0,$$

and, consequently the vector of polarisation is directed along the electric axis. For extension along the X-axis $P = P_x = d_{11}t_{11}$, and along the Y-axis $P = P_x = -d_{11}t_{22}$. For a hydrostatic pressure $t_{11} = t_{22} = t_{33}$ and consequently there is no piezoelectric effect.

In the rhombic γ-modification of Rochelle salt, at temperatures well above the Curie point (+22·5° C) there are three mutually perpendicular non-polar diad axes of symmetry and these are simultaneously the optic and electric axes of the crystal.

In the crystallographic system of coordinates, the piezoelectric tensor has the form

	t_{11}	t_{22}	t_{33}	t_{23}	t_{31}	t_{12}
P_x	0	0	0	d_{14}	0	0
P_y	0	0	0	0	d_{25}	0
P_z	0	0	0	0	0	d_{36}

and is determined by three independent components. The piezoelectric polarisation is determined by the relations

$$P_x = d_{14}t_{23}; \quad P_y = d_{25}t_{31}; \quad P_z = d_{36}t_{12},$$

which show that the piezo-effect in any given case can only occur as a result of shear stresses. The piezoelectric modulus has the following values in the crystals we are considering: α-quartz (18° C) $d_{11} = 6{\cdot}8\times10^{-8}$, $d_{14} = -1{\cdot}8\times10^{-8}$ c.g.s. units; Rochelle salt (30° C) $d_{14} = 2{\cdot}0\times10^{-5}$, $d_{25} = -1{\cdot}65\times10^{-8}$, $d_{36} = 35{\cdot}4\times10^{-8}$ c.g.s. units. One of the piezoelectric moduli of Rochelle salt has a very large value; the corresponding component of polarisation is directed along the axis which is coincident with the ferroelectric axis of the crystal.

4. Ferro- and Antiferroelectrics

The Basic Properties of Ferroelectrics

Investigations of the dielectric properties of certain substances show an 'anomalous' dependence of the dielectric constant on the strength of the electric field. For example, we may observe a breakdown in the simple linear relation $Q = CV$, which relates the charge to the potential difference between the plates of a condenser when the space between them is filled with the dielectric. Thus, for Rochelle salt it is found that for small V, the charge increases approximately proportional to V^2. As V increases, the increase of ε becomes slower and the curve of $\varepsilon(V)$ passes through a maximum (Fig. 129*a*): $\varepsilon_m = 9500$. In the absence of an external field, the dielectric constant remains at a large value of the order of several thousands. Investigations at varying temperatures show the presence of sharp maxima on the curve of $\varepsilon = \varepsilon(T)$ (Fig. 129*b*); these indicate phase transformations connected with large changes in the dielectric properties. The temperature maxima T_C on the curve $\varepsilon(T)$ are called the dielectric Curie points. In Rochelle salt two Curie points are found, the higher at $T_1 = +22{\cdot}5°$ C and the lower at $T_2 = -15°$ C.

When a crystal is investigated in a varying electric field between the Curie points, an electric hysteresis loop is observed (Fig. 130). When the field strength is increased the polarisation at first increases along the curve $OABC$ and tends to saturation. When the field is reduced to $E = 0$ the dielectric remains in the polarised state. The residual polarisation P_s is really OD, but it is usually measured by the ordinate OE, obtained by extrapolating the linear part of the curve BC. To destroy the *spontaneous polarisation* it is necessary to apply an electric field with reversed direction, $OF = E_c$, where E_c is called the *coercive force*. The existence of spontaneous polarisation and the possibility of reversing its direction by means of applying a field which exceeds the coercive force in absolute magnitude, are characteristics of the ferroelectric state

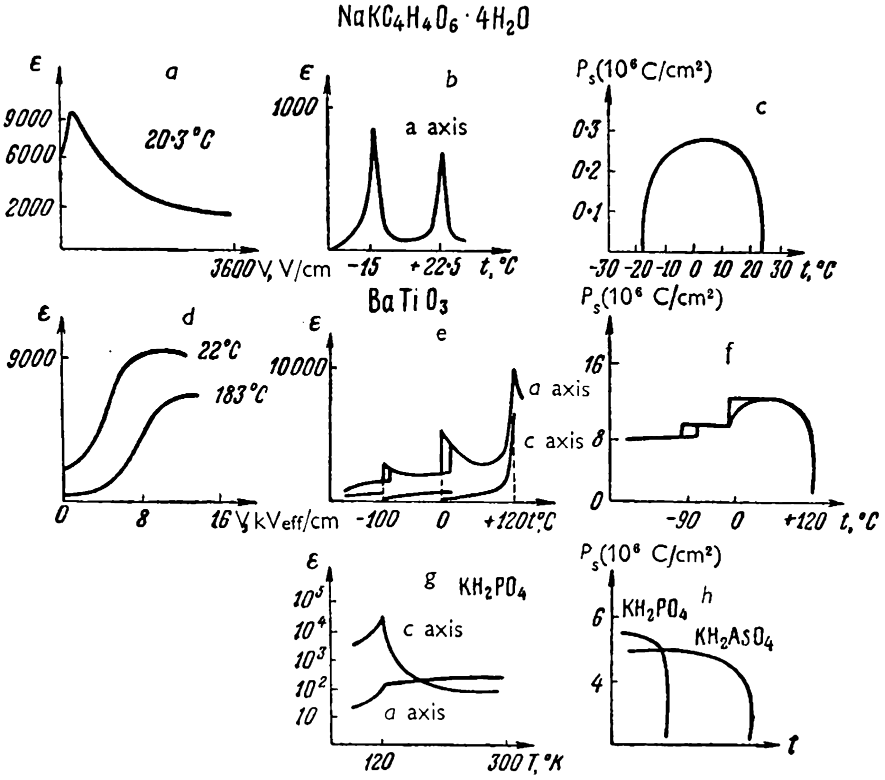


FIG. 129. Dependence of the dielectric constant on the field strength ε(V) (curves *a*-*d*) and on temperature ε(*T*) (curves *b*, *e*, *g*) and the dependence of the spontaneous polarisation on temperature $P_s(T)$ (curves *c*, *f* and *h*) for Rochelle salt, barium titanate and potassium phosphate.

of a substance. At temperatures above T_2 and below T_1, the hysteresis loop is absent and there is a linear dependence of P on E, corresponding to the paraelectric state of Rochelle salt. In the ferroelectric range the magnitude of the spontaneous polarisation depends on temperature (Fig. 129*b*). The phase transformation at the dielectric Curie points is accompanied by changes in the symmetry and dimensions of the unit cell, changes in the energy of the crystal, the piezoelectric moduli and so on. Thus the paraelectric modification belongs to the rhombic system and the ferroelectric modification to the monoclinic system, with the ferroelectric axis coincident with the *a*-axis of the rhombic modification. In Rochelle salt, a second-order phase transformation takes place but first-order phase transformations are possible in other ferroelectrics. Substitution of deuterium for hydrogen in water leads to a marked broadening of the ferroelectric range. As Kurchatov and Eremeev showed some time ago, isomorphic transformations have a marked influence on the Curie

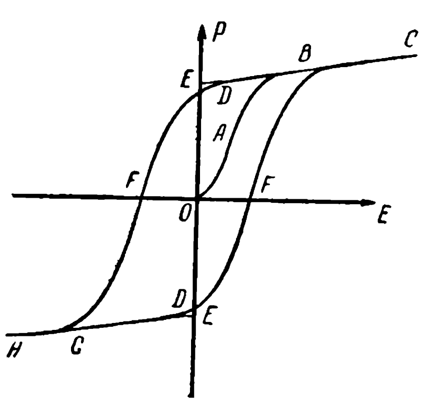


FIG. 130. A hysteresis loop.

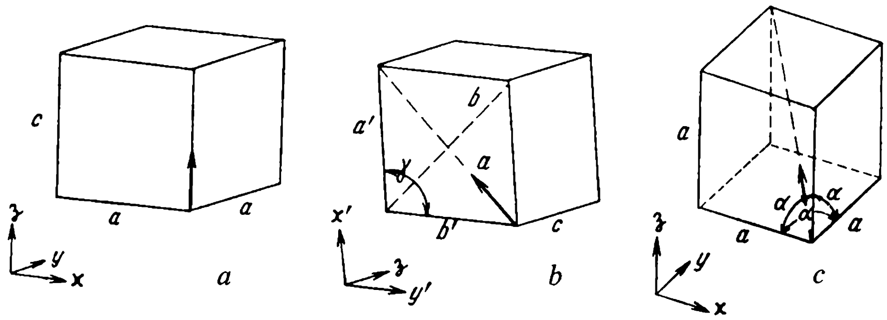


FIG. 131. The unit cells of the three ferroelectric modifications of $BaTiO_3$: (*a*) tetragonal, (*b*) rhombic (in the monoclinic orientation this shows the connection with the initial cubic structure), (*c*) rhombohedral.

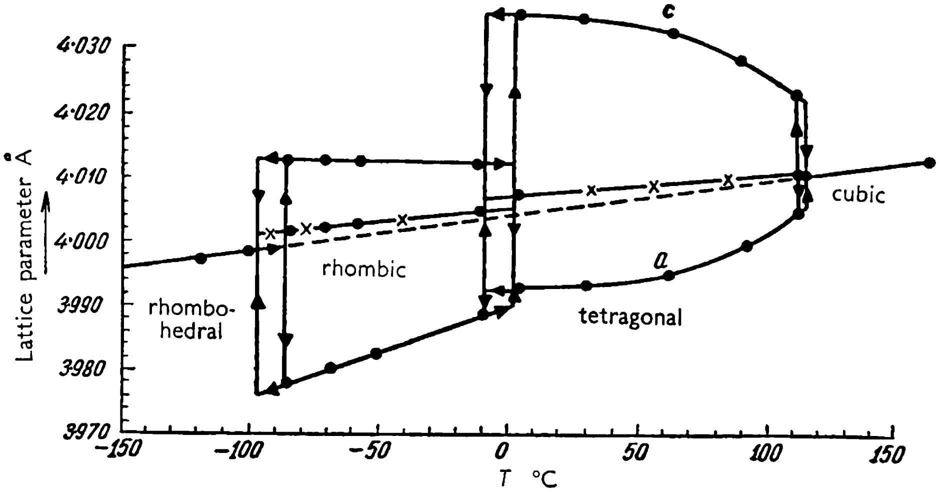


FIG. 132. Change of the lattice parameters of the various modifications of $BaTiO_3$ as a function of temperature.

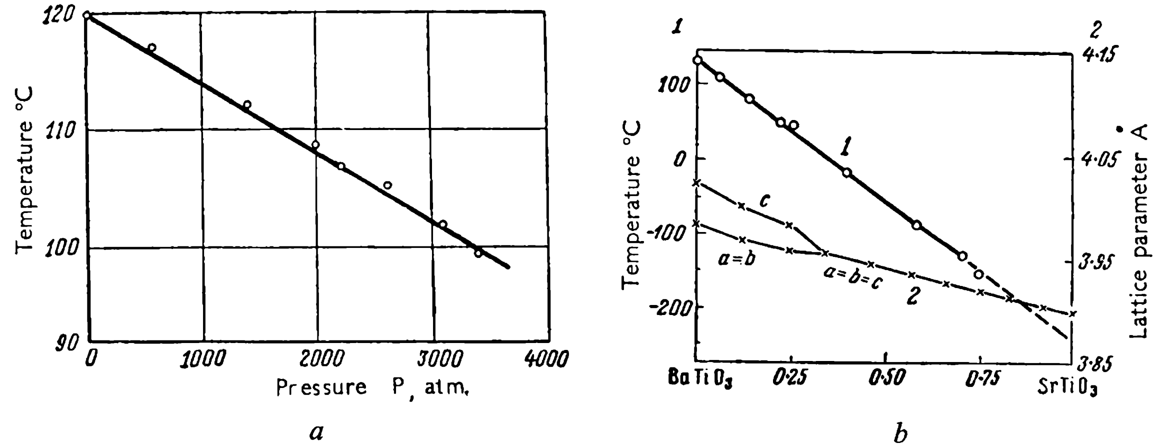


FIG. 133. The dependence of the Curie temperature: (*a*) of a single crystal of $BaTiO_3$ on the hydrostatic pressure; (*b*) of polycrystalline solid solutions (Ba, Sr) TiO_3 on concentration (curve 1). Curve 2 shows the values of the lattice parameters at room temperature.

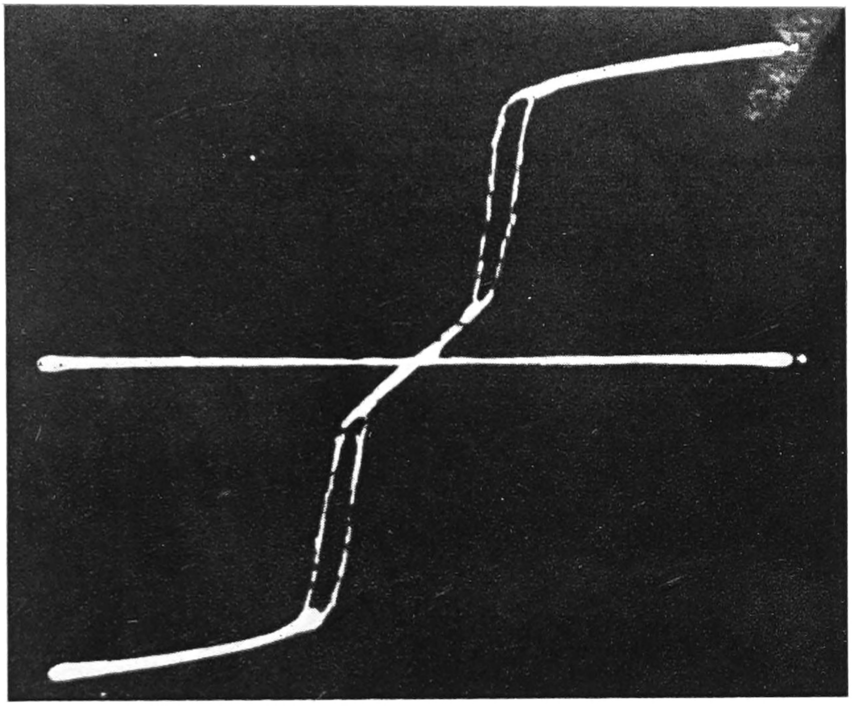


FIG. 134. Double hysteresis loop in $BaTiO_3$ above the Curie point.

temperature. Thus the addition of 1% of an ammonium salt to Rochelle salt narrows the interval $T_2 - T_1$ by a factor of 2, while with 12% of ammonium salt there is only one Curie point.

The next group of ferroelectrics—the phosphates and arsenates of the alkali metals—was discovered in 1935. Here the transition to the ferroelectric state takes place at low temperatures (Table 55). Fig. 129*g*, *h* show the change in the properties of potassium phosphate KH_2PO_4 in this transition, and from them we see that the spontaneous polarisation and ε_m are much greater than for Rochelle salt. The crystal symmetry changes from tetragonal to rhombic.

Large values for the ferroelectric properties were discovered by Vul (1945) for barium titanate and for a series of other substances discovered about that time and having the perovskite type of structure. These substances are double oxides of metals with the simple composition ABO_3 and are stable, high-melting point compounds with high Curie points which makes them potentially more useful. The perovskite structure is very simple being cubic in the paraelectric regions. Changes of the electric properties of barium titanate are shown in Fig. 129*d*, *e*, *f*. In barium titanate there are three ferroelectric transitions at $T_1 = 120°$ C; $T_2 = 5°$ C and $T_3 = -80°$ C (Fig. 129*e*). The ferroelectric modifications have respectively tetragonal, rhombic and rhombohedral symmetry and the unit cells of all of them are obtained by small deformations starting from a cubic cell and corresponding to an extension along the various axes of symmetry which coincide with the directions of the ferroelectric axes (Fig. 131). Experiments on single crystals have shown that the spontaneous polarisation changes discontinuously at the Curie points. At the same time the dimensions of the unit cell change discontinuously (Fig. 132). All this indicates first order phase transformations.

With increasing external pressure, the Curie temperature falls linearly (Fig. 133*a*). There is also a linear change of the Curie temperature as a function of concentration in the solid solution $BaTiO_3$—$SrTiO_3$ (Fig. 133*b*). As can be seen from the curve of lattice parameters, replacing the barium ions by the smaller ions of strontium leads to a compression of the unit cell of the solid solution and this is accompanied by a lowering of the Curie temperature.

An external electric field assists polarisation and raises the Curie temperature of a ferroelectric. If we apply a sufficiently large electric field to barium titanate, somewhat above its Curie temperature (120° C), that is in the paraelectric range, we observe a sudden increase in the temperature of the substance by 0·5°. This electrothermal effect can be explained if it is assumed that the field produces a ferroelectric transition and that the latent heat of the transition $q = c\Delta T$ contributes heat to the specimen. On the basis of the Neumann-Kopp Law we have, that for $BaTiO_3$ $c = 30$ cal/mol. °K. Hence $q = 15$ cal/mole.

A forced ferroelectric transition can be demonstrated visually on the screen of a cathode ray oscilloscope. In a weak field, the crystal behaves as a paraelectric and gives a linear dependence of P on E while, when a critical value E_c is reached, the crystal transforms into the ferroelectric state and there is a sharp increase in the polarisation. Thus if a strong alternating field is applied to a single crystal of barium titanate above the Curie point we obtain a double hysteresis loop (Fig. 134) while if the field is further increased the polarisation tends to saturation. Below the Curie temperature we obtain a hysteresis loop the same as in an ordinary ferroelectric. The ferroelectric transition can take place in the electric field within a narrow temperature region, the limiting value being 10° above the Curie point. The existence of a critical field different from zero, $E_c \neq 0$ is a strong argument for a first order phase transition

Table 55 Crystals with Ferroelectric Phases

Common name and formula	Symmetry of the paraelectric phase	Ferro-electric axis	Space group		T_c °C	P_s, $\mu C/\mathrm{cm}^2$	E_c, V/cm	$\varepsilon_{\max}$	Refer-ence	Notes
			Para-electric	Ferro-electric						
Rochelle salt $NaKC_4H_4O_6 \cdot 4H_2O$	rhombic $a = 11{\cdot}89$; $b = 14{\cdot}26$; $c = 6{\cdot}22$	[100]	$P2_12_12_1$	$P2_1$	24 −18	0·27	200	~9,000		
Deuterated Rochelle Salt $NaKC_4H_2D_2O_6 \cdot 4D_2O$			$P2_12_12_1$	$P2_1$	35 −24					
Lithium ammonium tartrate monohydrate $Li(NH_4)C_4H_4O_6 \cdot H_2O$			$P2_12_12_1$		−167					
Potassium dihydrogen phosphate KH_2PO_4	tetragonal $a=b=7{\cdot}45$; $c = 6{\cdot}96$	[001]	$D_{2d}^{12}-F\bar{4}d2$	$C_{2v}^{19}-Fdd$	−140	5·3		~10^5		
Potassium deuterophosphate KD_2PO_4			$D_{2d}^{12}-F\bar{4}d2$	$C_{2v}^{19}-Fdd$	−60					
Potassium arsenate KH_2AsO_4			$D_{2d}^{12}-F\bar{4}d2$	$C_{2v}^{19}-Fdd$	−180					
Rubidium phosphate RbH_2PO_4			$D_{2d}^{12}-F\bar{4}d2$	$C_{2v}^{19}-Fdd$	−123					
Barium titanate (Perovskite type) $BaTiO_3$	cubic $a = 4{\cdot}00$	[001]	$O_h^1-P_m3m$		118	26	500	~10,000		
Potassium niobate (Perovskite type) $KNbO_3$			$O_h^1-P_m3m$		435	26				
Potassium tantalate (Perovskite type) $KTaO_3$			$O_h^1-P_m3m$		−260					
Lead titanate (Perovskite type) $PbTiO_3$			$O_h^1-P_m3m$							

Lithium niobate (Ilmenite type) $LiNbO_3$	rhombo-hedral a = 5·49; α = 55° 53′;	[111]	trigonal	C_{3v}^6-R3c						
Lithium tantalate (Ilmenite type) $LiTaO_3$		[111]								
Guanidine aluminium sulphate $[C(NH_2)_3]\,[Al(H_2O)_6]\,[SO_4]_2$	hexagonal	[111]	$Pa3$	C_{3v}^2-P3im	200	0·35 (20° C)	$1-3\cdot10^3$ 1,500 (18° C)	6	1–3	Decomposes, losing its water of crystallization below T_c
Guanidine aluminium selenate $[C(NH_2)_3]\,[Al(H_2O_6)]\,[SeO_4]_2$		[111]		C_{3v}^2-P3im						
Guanidine aluminium alum $[C(NH_2)_3(H_2O)_6]\cdot$ $\cdot\,[Al(H_2O)_6]\,[SO_4]_2$	cubic	[111]	T^4-P2_13	$C_2^2-P2_1$	~-80				4	
Methyl ammonium aluminium alum $[(CH_3NH_3)(H_2O)_6]\cdot$ $\cdot\,[Al(H_2O)_6]\,[SO_4]_2$		[111]	T^4-P2_13	$C_2^2-P2_1$	−98	0·6	15,000	ε_{111} = 60		
Urea chrome alum $[(Co(NH_3)_2)(H_2O)_6]\cdot$ $\cdot\,Cr(H_2O)_6]\,[SO_4]_2$		[111]	T^4-P2_13	$C_2^2-P2_1$	−113	0·1				
Thiourea $CS(NH_2)_2$	rhombic	[010]	$D_{2h}^{16}-Pnma$		−105	3·1	< 1,000		5	
Ammonium sulphate $(NH_4)_2SO_4$	rhombic a = 5·97; b = 10·60; c = 7·76;	[100]	$D_{2h}^{16}-Pnma$		−49·5	0·25	2,000		6	
Ammonium cadium sulphate $(NH_4)_2Cd_2(SO_4)_3$	cubic a = 10·35	[111]	cubic		−186	0·3	15,000	ε_{111}=10·2	7	$K_2Mg_2(SO_4)_3$

Common name and formula	Symmetry of the paraelectric phase	Ferro-electric axis	Space group		T_c °C	P_s, $\mu C/cm^2$	E_c, V/cm	ε_{max}	Refer-ence	Notes
			Para-electric	Ferro-electric						
Triglycine sulphate $(NH_2CH_2COOH)_3H_2SO_4$	monoclinic $a = 9{\cdot}15$; $b = 12{\cdot}69$; $c = 5{\cdot}73$; $\beta = 106°$	[010]	?	$C_2^2 - P2_1$	47	2·2 (20°)	220		8, 9, 10	Diglycol sulphate is rhombic and paraelectric
Triglycine selenate $(NH_2CH_2COOH)_3H_2SeO_4$					22					
Triglycinefluoroberylliate $(NH_2CH_2COOH)_3H_2BeF_4$					70	3·2	5,000	$\varepsilon \sim 5{,}000$	11	
Silver glycinitrate $(NH_2CH_2COOH)AgNO_3$	monoclinic $\beta = 100°$; $a = 5{\cdot}53$; $b = 19{\cdot}58$; $c = 5{\cdot}51$	[010]	$P2_1/a$		−55	0·55	740	$\varepsilon_6 = 650$	11	
Dicalcium strontium propionate $Ca_2S_r(CH_3CH_2COO)_6$	tetragonal	[001]			8·5	0·3	13,000		13	Isomorphic compounds with strontium also exist
Manganese diglycine chloride dihydride $(NH_2CH_2COOH)_2 \cdot$ $\cdot MnCl_2 \cdot 2H_2O$									14	
Diglyclinitrate $(NH_2CH_2COOH)_2 \cdot HNO_3$	monoclinic $a = 9{\cdot}50$; $b = 5{\cdot}11$; $c = 9{\cdot}35$; $\beta = 98{\cdot}8°$; $z = 2$	[101]	$P2_1/a$	P_a	−67	0·6	400	600	15	

Table 55—REFERENCES

1. A. N. Holden, B. T. Matthias, W. J. Merz and J. P. Remeika: *Phys. Rev.*, **98**, 546 (1955); **101**, 962, (1956).
2. Z. I. Ezhkova, G. S. Zhdanov and M. M. Umansky: *Kristallografiya* **3**, 231 (1958).
3. L. A. Varfolomeeva, G. S. Zhdanov and M. M. Umansky: *Kristallografiya* **3**, 368 (1958).
4. R. Pepinsky, F. Jona and G. Shirane: *Phys. Rev.*, **102**, 1181 (1956).
5. A. L. Solomon: *Phys. Rev.*, **104**, 1191 (1956).
6. B. T. Matthias and J. P. Remeika: *Phys. Rev.*, **103**, 262 (1956).
7. P. Iona and R. Pepinsky: *Phys. Rev.*, **103**, 1126 (1956).
8. B. T. Matthias, C. E. Miller and J. P. Remeika: *Phys. Rev.*, **104**, 849 (1956).
9. E. A. Wood and A. N. Holden: *Acta Cryst.*, **10**, 145 (1957).
10. Z. I. Ezhkova, G. S. Zhdanov and M. M. Umansky: *Kristallografiya* **4**, 249 (1959).
11. S. Hoshino, T. Mitsui, F. Jona and R. Pepinsky: *Phys. Rev.*, **107**, 1255 (1957).
12. R. Pepinsky, Y. Okaya, D. P. Eastman and T. Mitsui: *Phys. Rev.*, **107**, 1538 (1957).
13. B. T. Matthias and J. P. Remeika: *Phys. Rev.*, **107**, 1727 (1957).
14. R. Pepinsky, K. Vedam and Y. Okaya: *Phys. Rev.*, **110**, 1309 (1958).
15. R. Pepinsky, K. Vedam, S. Hoshino and Y. Okaya: *Phys. Rev.*, **111**, 430 (1958).

and indeed if the discontinuity is not large it is often difficult to distinguish first order transitions from second order transitions by other methods.

Changes of the dielectric constant above the Curie point in the paraelectric region are described by a relation analogous to the Curie-Weiss Law in ferromagnetism

$$\varepsilon = \varepsilon_0 + \frac{C}{T - T_0} \qquad \ldots [9.48]$$

Here T_0 is the Curie temperature defined above, which agrees with the transition temperature T_e for second order transitions but is less than T_c for first order transitions.

New Ferroelectrics

Within recent years, investigations of ferroelectricity have led to the discovery of a large number of new ferroelectrics and have shown that ferroelectricity is a more common phenomenon than had been thought previously. Basic data on the composition and properties of ferroelectric substances are given in Table 55, where amongst the perovskite group we could have included tungsten trioxide, the ideal structure of which (ReO_3 type) is obtained from the perovskite structure ABO_3 by taking away the cations of type A. Lithium niobate and tantalate have a structure of the ilmenite or corundum type. Alums with complex cations have a cubic structure in their paraelectric modification. Groups of guanidine-aluminium sulphates and groups of triglycine sulphates which have their Curie points above room temperature may have practical importance.

The majority of the known perovskite structures behave differently from $BaTiO_3$; $PbTiO_3$ and $KTaO_3$ have only one ferroelectric transition and potassium tantalate has a low Curie point ($T_c = 13°$ K)—so low that it is difficult to reconcile with the fact that the establishment of the ferroelectric state requires ionic movements. Again, it is curious that an isomorphic crystal $NaTaO_3$ has a very high Curie point ($T_c = 475°$ C). In compounds of the perovskite groups ($PbZrO_3$, $PbHfO_3$) and also in tungsten trioxide the transition diagrams become more complex, this being connected with the appearance of antiferroelectricity. In $NaNbO_3$ four phase transitions have been established (see antiferroelectrics below).

Only potassium niobate $KNbO_3$ behaves in the same way as barium titanate having three ferroelectric transitions with Curie points $T_1 = 415°\text{ C}$, $T_2 = 210°\ C$, $T_3 = -40°\text{ C}$ each somewhat higher than the corresponding points in $BaTiO_3$. The symmetry and the order in which these phases occur as the temperature is lowered in the same way as in barium titanate.

The Thermodynamics of the Ferroelectric Phase Transitions

(*a*) **Free Energy of a Crystal with one Ferroelectric Axis.** We may take the polarisation P and the field strength E as thermodynamic parameters which determine the state of a dielectric in an electric field. They are connected with the free energy F by the relation

$$E = \frac{\partial F}{\partial P}. \qquad \ldots [9.49]$$

For the majority of dielectrics we have a linear relation

$$P = \eta E, \qquad \ldots [9.50]$$

where η is the dielectric constant. The free energy of a one-dimensional dielectric depends parabolically on the polarisation

$$F = \tfrac{1}{2}\alpha P^2, \qquad \ldots [9.51]$$

where

$$\alpha = \frac{1}{\eta} = \frac{4\pi}{\varepsilon - 1}. \qquad \ldots [9.52]$$

With ferroelectrics we have a non-linear dependence of the polarisation on the field strength and we have to take higher terms to obtain the free energy. However ferroelectric transitions are accompanied by small changes of energy and structure of the crystals, so, as a first approximation, we may assume that the same expression for the free energy applies to the paraelectric and ferroelectric phases.

For a crystal with one ferroelectric axis the polarisation vector can have only two orientations which coincide with the positive and negative directions of that axis. The energy of the crystal for the two directions of polarisation $+P$ and $-P$ must be the same. Thus to obtain the free energy as a function of the degree of polarisation we require only even terms with

$$F = \tfrac{1}{2}\alpha P^2 + \tfrac{1}{4}\beta P^4 + \tfrac{1}{6}\gamma P^6 + \ldots \qquad \ldots [9.53]$$

For the field strength we have

$$E = \frac{\partial F}{\partial P} = \alpha P + \beta P^3 + \gamma P^5 + \ldots \qquad \ldots [9.54]$$

The coefficients $\alpha, \beta, \gamma \ldots$ depend on the basic thermodynamic parameters: temperature, pressure, volume and field strength. In the majority of experiments with ferroelectrics the external pressure does not change (free crystal) so we can consider an isobaric process ($p = \text{const.}$).

Near the Curie point the dielectric constant changes rapidly with temperature according to the Curie-Weiss Law and reaches very high values ($\varepsilon \gg 1$). In this case we obtain from [*52*] and [*48*]

$$\alpha \approx \frac{4\pi}{\varepsilon} = \frac{4\pi}{C}(T-T_0). \qquad \ldots [9.55]$$

The coefficient α depends linearly on temperature and changes sign at temperature $T = T_0$.

(*b*) **First and Second Order Ferroelectric Transitions.** Depending on the number, signs and magnitudes of the coefficients of [*53*] the curve of F takes essentially different forms. If the signs of all the terms are positive ($\alpha, \beta, \gamma \ldots > 0$), then the curve has one minimum at the origin $P = 0$ (Fig. 135*a*) and the paraelectric state is stable in all ranges of the variable P.

If however, one of the coefficients in the expression is negative, then in the absence of an external field ($E = 0$), equation [*54*] becomes $dF/dP = 0$, and can have other solutions than $P = 0$. For these $P \neq 0$ and they correspond to a state of spontaneous polarisation of the crystal. Whether or not these states are stable relative to the paraelectric state depends on the relationship between the free energies of these states.

Case 1. *Second order phase transition.* If we take two terms of the series and suppose that $\alpha < 0$, $\beta > 0$, then the curve of F has a maximum at the origin and two symmetrically placed minima B and B' (Fig. 135*b*). The state with spontaneous polarisation is stable and the minima of the F curve describe a ferroelectric state if the height of the potential barrier is not too large and transitions can take place between B and B' in relatively weak fields. If the coefficient α changes sign [*55*] at $T = T_0$, then at that temperature there is a ferroelectric transition. If $\alpha = \alpha(T)$ is a continuous function, then the spontaneous polarisation $P = P(T)$ is also a continuous function of temperature (Fig. 136*a*) reducing to zero at $T = T_0$: $P(T_0) = 0$. The energy of the crystal changes continuously and the transition takes place with no latent heat ($q = 0$) but with a discontinuity in specific heat. This is typical of second order phase transitions.

Case 2. *First order phase transition.* If we take three terms in the series and assume that $\alpha > 0$, $\beta < 0$, then the curve of F has three minima; one at the origin and two B and B' symmetrically placed on either side of it (Fig. 135*b*). If the relationship between the depths of the maxima depends on temperature and the potential barriers are not too large, there will be a transition from the para- to the ferro-electric state. Unlike the preceding case, the spontaneous polarisation (Fig. 136*b*) and the energy of the crystal at the transition point both have a discontinuity. That is, the latent heat of transition has a finite value ($q \neq 0$). This is typical of first order phase transitions.

(*c*) **The Condition for Equilibrium.** The turning points of the free energy are determined by the equation

$$\frac{dF}{dP} = 0,$$

which, as equation [*54*] shows breaks up into two equations. One of them $P = 0$ gives a turning point at the origin and the other

$$\alpha + \beta P^2 + \gamma P^4 + \ldots = 0 \qquad \ldots [9.56]$$

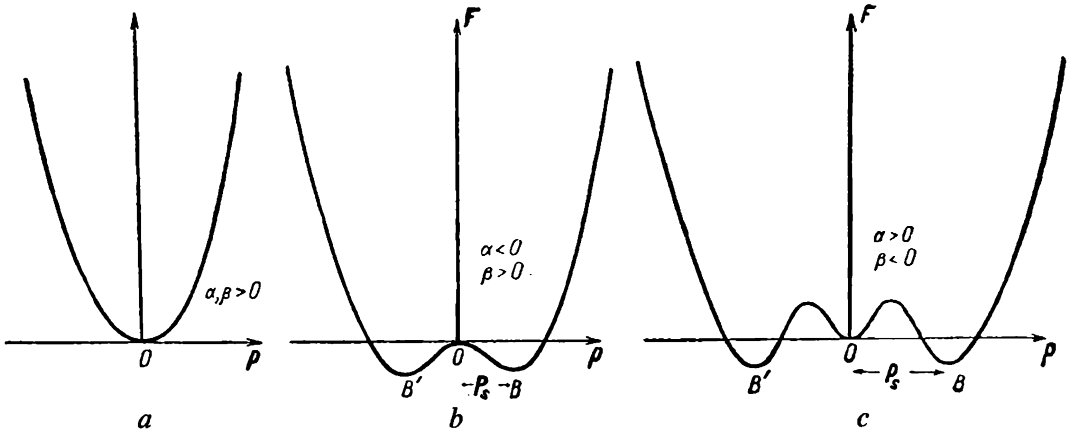


FIG. 135. The shape of the free energy curves: (*a*) paraelectric phase; (*b*) ferroelectric phase (second order transition), (*c*) ferroelectric phase (first order transition).

gives turning points at values of P differing from zero. We can now find the solutions of this equation for the cases analysed graphically above.

Case 1. Equation (56) takes the form

$$\alpha+\beta P^2 = 0,$$

and from this we obtain the value of the spontaneous polarisation

$$P^2 = -\frac{\alpha}{\beta}. \qquad \ldots [9.57]$$

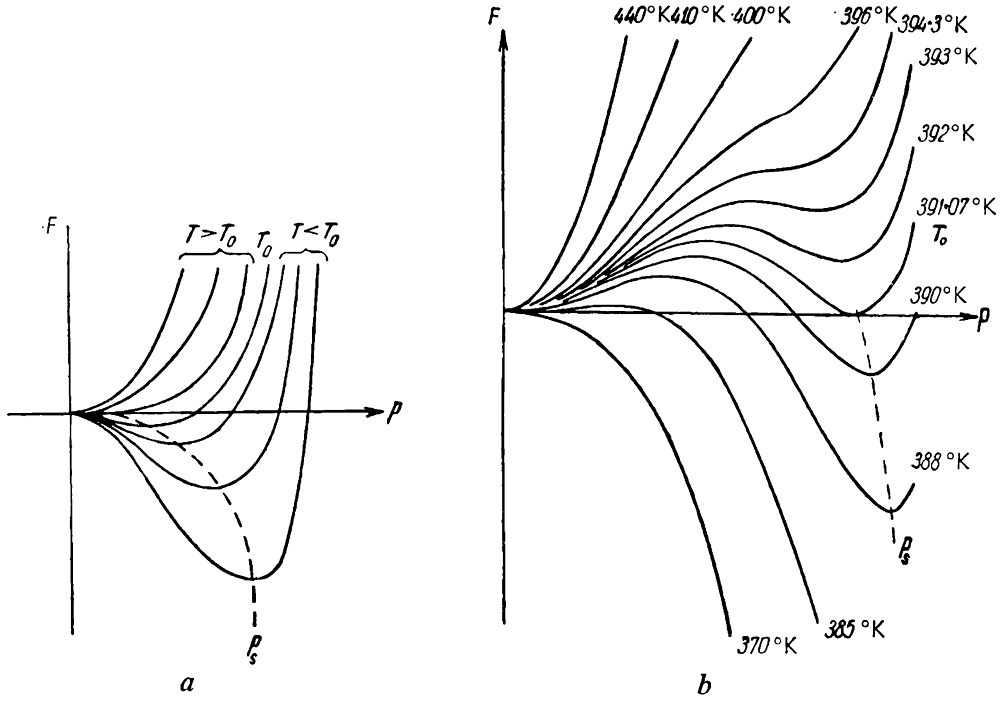


FIG. 136. Free energy curves of a free crystal at various temperatures. Ferroelectric transition (*a*) second order (Rochelle salt KH_2PO_4), (*b*) first order ($BaTiO_3$).

Since $\beta > 0$, this solution is physically possible in the region where α is negative. If for α we assume a linear dependence [*55*] on temperature and take β to be independent of temperature, we obtain a parabolic law for the spontaneous polarisation

$$P^2 = c(\theta - T), \qquad \ldots [9.58]$$

where c is some constant.

Case 2. Equation (56) takes the form

$$\alpha + \beta P^2 + \gamma P^4 = 0. \qquad \ldots [9.59]$$

Since there is a minimum at the origin, the equilibrium conditions for the para- and ferroelectric phases are determined by equating their free energies. This leads to a second equation of the form

$$\alpha + \tfrac{1}{2}\beta P^2 + \tfrac{1}{3}\gamma P^4 = 0, \qquad \ldots [9.60]$$

which is obtained from [*53*] by putting $F = 0$. Now equations [*59*] and [*60*] must be valid simultaneously and this imposes limits on the coefficients of the equation. Subtracting one equation from the other we obtain the value of the spontaneous polarisation at the transition point

$$P_0^2 = -\tfrac{3}{4}\frac{\beta_0}{\gamma_0}.$$

Putting this value into one or the other of equations [*59*] and [*60*] we obtain a relationship between the coefficients in the series for the free energy

$$\alpha_0\gamma_0 = \tfrac{3}{16}\beta_0^2. \qquad \ldots [9.61]$$

Equation [*61*] must be satisfied for a first order ferroelectric transition. At the transition point, the spontaneous polarisation changes discontinuously from $P = 0$ to

$$P_0^2 = -4\frac{\alpha_0}{\beta_0} = -\tfrac{3}{4}\frac{\beta_0}{\gamma_0}. \qquad \ldots [9.62]$$

Since $\gamma > 0$ we must have $\beta < 0$ and $\alpha > 0$.

The entropy of the system is given by the expression

$$S = \left(\frac{\partial F}{\partial T}\right)_{P,\sigma} \approx -\tfrac{1}{2}P^2\frac{\partial\alpha}{\partial T}, \qquad \ldots [9.63]$$

from which we obtain the change of entropy

$$\Delta S = S - S_0 \approx -\tfrac{1}{2}P_s^2\left(\frac{4\pi}{C}\right). \qquad \ldots [9.64]$$

This formula is valid for both types of transition. For a first order transition $\Delta S = q$, where q is the latent heat and $P_s = P_0$.

The position of the Curie point depends on the values of other thermodynamic parameters. However the shift of the Curie point produced by an external electric field is given by the formula

$$\frac{\partial T_k}{\partial E} = -\frac{\Delta P}{\Delta S} = -\frac{P_0}{\Delta S} = \frac{C}{4\pi}\left(\frac{2}{P_0}\right) = \left(\frac{C}{4\pi}\right)\left(-\frac{\beta_0}{\alpha_0}\right)^{\frac{1}{2}}. \qquad \ldots [9.65]$$

The shift produced by external pressure can be determined from the Clausius-Clapeyron Equation

$$\frac{\partial T_k}{\partial p} = \frac{\Delta V}{\Delta S}, \qquad \ldots [9.66]$$

Where ΔV is the (discontinuous) change of volume produced by heating at the transition point.

(*d*) **Comparison with Experiment for Barium Titanate.** Using [*61*] and [*62*] we can express the values of the coefficients β_0 and γ_0 at the Curie point in terms of measurable quantities—the spontaneous polarisation P_0 and α_0:

$$\beta_0 = -4\frac{\alpha_0}{P_0^2}; \quad \gamma_0 = 3\frac{\alpha_0}{P_0^4}, \qquad \ldots [9.67]$$

where

$$\alpha_0 \approx \frac{4\pi}{C}(T_c - T_0). \qquad \ldots [9.68]$$

For barium titanate, the difference between the Curie temperature T_c and the Curie parameter T_0 is $T_c - T_0 = 11°\text{ C}$. In this case we obtain from the thermodynamic relation [*67*]

$$\beta_0 = -10{\cdot}8\times10^{-13}; \quad \gamma_0 = 28{\cdot}2\times10^{-23} \text{ c.g.s. units.} \qquad \ldots [9.69]$$

The shift of the Curie point in an electric field as measured by means of a double hysteresis loop is

$$\beta_0 = -6{\cdot}8\times10^{-13}; \quad \gamma_0 = 22{\cdot}8\times10^{-23} \text{ c.g.s. units.}$$

Recently, the dependence of the dielectric constant of barium titanate on polarisation has been measured, and compared with the value obtained from the series:

$$\frac{\partial^2 F}{\partial P^2} = \frac{\partial E}{\partial P} \approx \frac{4\pi}{\varepsilon} = \alpha + 3\beta P^2 + 5\gamma P^4. \qquad \ldots [9.70]$$

In these experiments the polarisation P was produced by a field of low frequency (60~) and varying amplitude, while the value of ε was found from the capacity of a condenser measured in a weak field of high frequency. At 120° C the measured value of β_0 was $-10{\cdot}0\times10^{-13}$ c.g.s. units which is in good agreement with the thermodynamic value. The linear temperature dependence of β in the range 120-150° C was also established. As temperature falls, the numerical value of β falls rapidly and if the linear dependence is maintained must pass through zero at $-175°$ C.

From the thermodynamic formula [*64*] we obtain for the entropy discontinuity

$$\Delta S = 0{\cdot}12 \text{ cal/mol. °C},$$

which agrees with the experimental value. It follows from formula [*65*] that the shift of the Curie point in an electric field is

$$\frac{\partial T_c}{\partial E} = 1{\cdot}6\times10^{-3} \text{ °C cm/V}. \qquad \ldots [9.71]$$

Measurements with ceramic specimens gave

$$\frac{\Delta T_c}{\Delta E} = 1{\cdot}2\times10^{-3} \text{ °C cm/V},$$

and measurements on single crystals of barium titanate gave $1{\cdot}4\times10^{-3}$ °C. cm/V. From [*66*] and using $\Delta V = -0{\cdot}06$ Å^3 per unit cell we obtain for the shift of the Curie point under the influence of pressure,

$$\frac{\partial T_c}{\partial p} = -6{\cdot}7\times10^{-9}\ {}^\circ\mathrm{C\ cm^2/dyne}, \qquad \ldots [9.72]$$

while the experimental value is

$$-5{\cdot}7\times10^{-9}\ {}^\circ\mathrm{C\ cm^2/dyne}.$$

These results show that the thermodynamic theory using empirical values of the coefficients of the series is in satisfactory agreement with experiment.

(*e*) **The Free Energy of a Cubic Ferroelectric.** In a cubic crystal there is not one but several equivalent ferroelectric axes. Let P_x, P_y and P_z be the components of the polarisation vector. Then we can write the free energy of a cubic crystal in the form

$$\begin{aligned} F &= \tfrac{1}{2}\alpha(P_x^2+P_y^2+P_z^2)+\tfrac{1}{4}\beta_1(P_x^4+P_y^4+P_z^4)+ \\ &+\tfrac{1}{2}\beta_2(P_y^2P_z^2+P_z^2P_x^2+P_x^2P_y^2)+\tfrac{1}{6}\gamma_1(P_x^6+P_y^6+P_z^6)+\ldots \end{aligned} \qquad \ldots [9.73]$$

The third term in this series expresses the interaction of the components of the polarisation along different crystal axes. If $\beta_2>0$, polarisation along one axis hinders the development of polarisation along an axis perpendicular to it.

Limiting the series to the terms in [*73*] we obtain the components of the field strength along the various axes of the crystal. For the *X*-axis we have

$$E_x = \frac{\partial F}{\partial P_x} = \alpha P_x+\beta_1P_x^3+\beta_2P_x(P_y^2+P_z^2)+\gamma_1P_x^5, \qquad \ldots [9.74]$$

with similar expressions for the two other axes. In the absence of an external field we have

$$E_x = \frac{\partial F}{\partial P_x} = 0, \quad E_y = \frac{\partial F}{\partial P_y} = 0, \quad E_z = \frac{\partial F}{\partial P_z} = 0, \qquad \ldots [9.75]$$

which leads to a system of equations for the turning values of the components of the crystal polarisation:

$$\left.\begin{aligned} P_x &= 0 \text{ or } \gamma_1P_x^4+\beta_1P_x^2+\beta_2(P_y^2+P_z^2)+\alpha = 0 \\ P_y &= 0 \text{ or } \gamma_1P_y^4+\beta_1P_y^2+\beta_2(P_z^2+P_x^2)+\alpha = 0 \\ P_z &= 0 \text{ or } \gamma_1P_z^4+\beta_1P_z^2+\beta_2(P_x^2+P_y^2)+\alpha = 0 \end{aligned}\right\}. \qquad \ldots [9.76]$$

Equation [*76*] satisfies the conditions of cyclic permutation and symmetry relative to P_x, P_y, P_z. On grounds of symmetry there are the following four systems of solutions of these equations:

$$\left.\begin{aligned} &P_x = P_y = P_z = 0 && (1) \\ &P_x = P_y = 0;\ \ \gamma_1P_z^4+\beta_1P_z^2+\alpha = 0 && (2) \\ &P_x = 0;\ \ P_y = P_z;\ \ \gamma_1P_z^4+(\beta_1+\beta_2)P_z^2+\alpha = 0 && (3) \\ &P_x = P_y = P_z;\ \ \gamma_1P_z^4+(\beta+2\beta_2)P_z^2+\alpha = 0 && (4) \end{aligned}\right\}. \qquad \ldots [9.77]$$

Solution (1) corresponds to the paraelectric phase; solutions (2), (3) and (4) are the cases when the ferroelectric axis is directed respectively along the edge, the face diagonal and the body diagonal of the cube, i.e. the tetragonal, the rhombic and the rhombohedral phases.

The values of the free energy corresponding to the solutions of [77] will be

$$\left.\begin{aligned} F &= 0 && (1)\\ F &= \tfrac{1}{6}\gamma_0 P_z^6+\tfrac{1}{4}\beta_1 P_z^4+\tfrac{1}{2}\alpha P_z^2 && (2)\\ F &= \tfrac{1}{3}\gamma_1 P_z^6+\tfrac{1}{2}(\beta_1+\beta_2) P_z^4+\alpha P_z^2 && (3)\\ F &= \tfrac{1}{2}\gamma_1 P_z^6+\tfrac{3}{4}(\beta_1+2\beta_2) P_z^4+\tfrac{3}{2}\alpha P_z^2 && (4)\end{aligned}\right\}. \qquad \ldots [9.78]$$

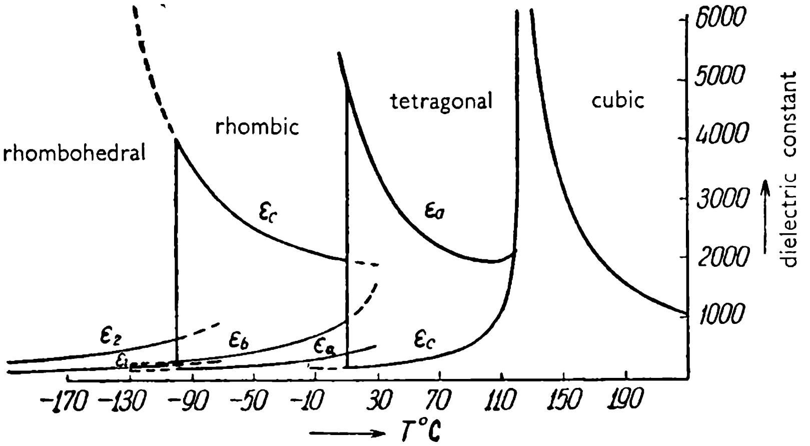


FIG. 137. Theoretical curves of the temperature dependence of the dielectric constant for a single crystal of barium titanate.

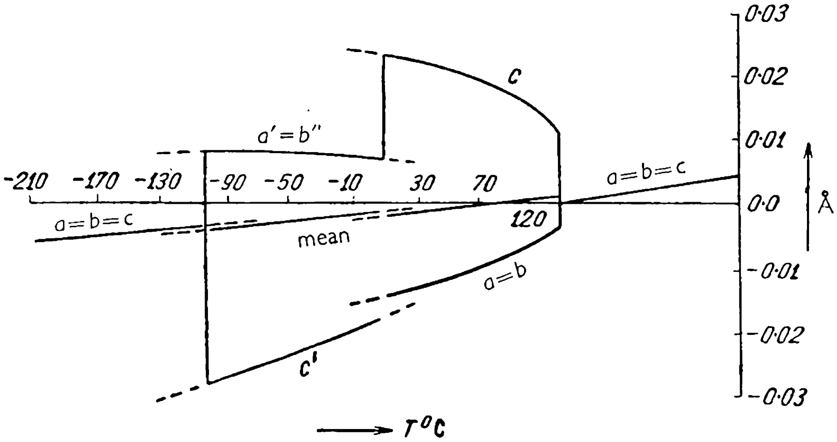


FIG. 138. Theoretical curves of the change of lattice parameters of various modifications of barium titanate.

The existence of these phases is possible if the conditions for a minimum F_0 are fulfilled. These are:

$$\left.\begin{aligned} &\alpha > 0 && && (1)\\ &\alpha+\beta_2 P_z^2 > 0, && \tfrac{1}{2}\beta_1+\gamma_1 P_z^2 > 0 && (2)\\ &\alpha+2\beta_2 P_z^2 > 0, && \tfrac{1}{2}\beta_1+\gamma_1 P_z^2 > 0 && (3)\\ & && \tfrac{1}{2}\beta_1+\gamma_1 P_z^2 > 0 && (4)\end{aligned}\right\}. \qquad \ldots [9.79]$$

If α is positive and sufficiently large, the minimum will be at $F = 0$ giving a stable paraelectric phase. Assuming that the coefficients β and γ are constants, then as α becomes smaller the minimum value of the free energy corresponds successively to solutions (2), (3) and (4), in that order. Thus the thermodynamic theory gives the number and succession of the ferroelectric phases in $BaTiO_3$. By use of four experimental quantities (T_c, T_0, C and P_s) the theory quantitatively describes the temperature dependence of the spontaneous polarisation, the dielectric constant (Fig. 137) and the changes of the cell parameters (Fig. 138) for all modifications of barium titanate.

(*f*) **Transition at Constant Volume.** An interesting result of the theory in this case is that $\beta_2 > 0$ for the first ferroelectric transition in barium titanate. Consequently in a compressed crystal there is a second order transition while in a free crystal there is a first order transition. The discontinuity in the latter case is a result of the form of the relation between polarisation and deformation. Fractional extension of the *c*-axis in the tetragonal transition is proportional to the square of the spontaneous polarisation: $\Delta c/c \sim P_s^2$.

The approximations in this theory arise from the use of the same coefficients in the series for the free energy of the various phases although the transition may be discontinuous. The success of such an approximation depends on the comparatively small discontinuities in ferroelectric transitions and so the theory agrees well with experiment except near the Curie point.

ANTIFERROELECTRICS

(*a*) **Basic Properties of Antiferroelectrics.** In some dielectrics there are phase changes which have almost all the characteristics of ferroelectric transitions: the dielectric constant passes through a maximum, there is a discontinuity in specific heat, there are changes in the symmetry and dimensions of the unit cell. However there are other essentially different characteristics: there is no hysteresis loop and no spontaneous polarisation. An example of such a dielectric is lead zirconate, $PbZrO_3$, which belongs to the perovskite family. As can be seen from Fig. 139 there is a sharp maximum on the $\varepsilon(T)$ curve at $T_c = 230°$ C. Above the maximum, the Curie-Weiss Law is obeyed with $C = 1{\cdot}6 \times 10^5$ °C, $T_0 = 193°$ C. At the transition point the symmetry of the crystal is lowered, and from cubic it becomes rhombic. The differences between the transition in lead zirconate and the ferroelectric transition in barium titanate are: (1) there is no hysteresis loop, (2) an electric field lowers the Curie point in $PbZrO_3$ (Fig. 140) but increases it in $BaTiO_3$, (3) in $PbZrO_3$ a superlattice arises based on a pseudotetragonal sub-cell a_t, c_t of the perovskite type with $a \approx \sqrt{2}a_t$, $b \approx 2\sqrt{2}a_t c \approx 2c_t$, (4) the tetragonal sub-cell is compressed $c_t/a_t \approx 0{\cdot}99$ whereas in $BaTiO_3$ there is an extension, (5) the volume of $PbZrO_3$ is reduced while the volume of $BaTiO_3$ is increased on cooling through the Curie point, (6) the refractive index of $PbZrO_3$ has its smallest value in the direction of the *a*-axis ($n_b - n_a = 0{\cdot}039$; $n_c - n_a = 0{\cdot}024$) and perpendicular to the pseudotetragonal axis, while in $BaTiO_3$ the refractive index has its smallest value in the direction of the tetragonal axis along which the movement of ions takes place.

The work which decided the nature of this transition was the solution of the structure of the low temperature modification of lead zirconate which was carried out on a single crystal by Savachugi *et al.* in 1951. This structure is shown in Fig. 141; here *a* and *b* are the directions of the axes of the rhombic structure. The multiplication in the volume of the unit cell comes about because below the transition point there are antiparallel displacements of the Pb ions, proceeding mainly on the plane *ab* and in a direction along the rhombic axis *a*; this explains the reduction in the ratio c_t/a_t. A superlattice is formed by the antiparallel arrangement of the double parallel chains; this explains the doubling of the *b*-axis as compared with the *a*-axis. As a result of antiparallel displacements in the lattice, or in the projection of the lattice, the centre of symmetry is retained. In centro-symmetric lattices neither electret nor ferroelectric states can arise.

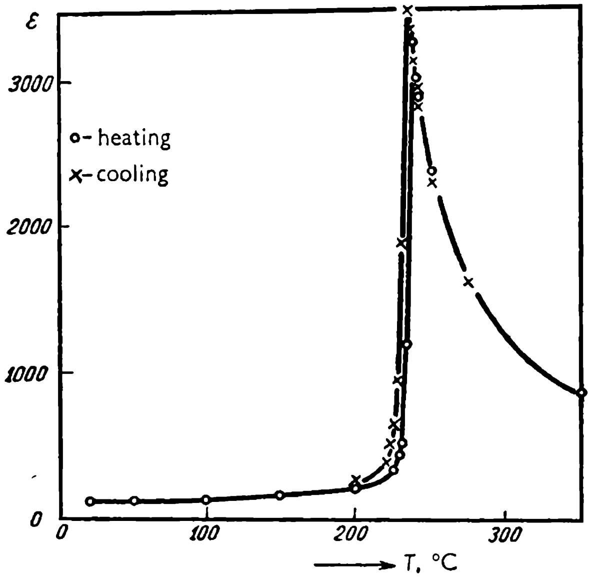


FIG. 139. Temperature dependence of the dielectric constant for a single crystal of PbZrO₃.

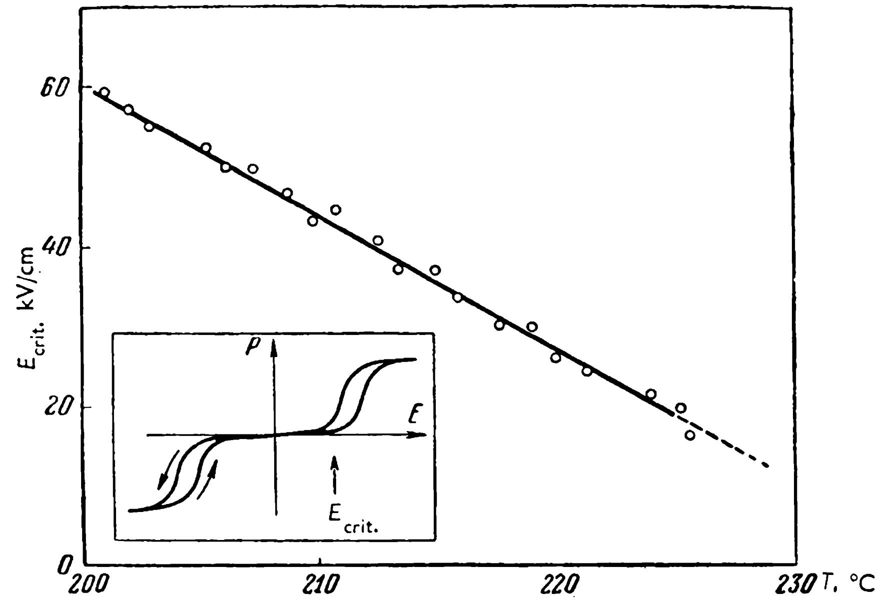


FIG. 140. Lowering of the transition point of $PbZrO_3$ in an electric field.

This explanation shows that an antiparallel arrangement of ions arises in the phase transition of lead zirconate and that this leads to mutual compensation of large elementary electric moments produced by the ionic displacements. The vectors of spontaneous polarisation of the sub-lattices of the crystal are directed antiparallel and mutually compensate each other. The resultant polarisation of the crystal is zero. Such a state of a dielectric is called *antiferroelectric* and the transition temperature is called the antiferroelectric Curie point.

It is worth pointing out that antiparallel arrangements of ions on some plane may be accompanied by parallel arrangements along its normal. This leads to the appearance of ferro- and antiferroelectric states in one and the same crystal. The ferroelectric effect may mask the antiferroelectric state. To settle the type and anisotropy of the state we need structural investigations and dielectric measurements on single crystals.

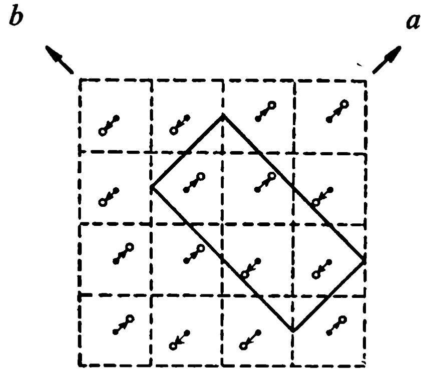


FIG. 141. Superlattice of $PbZrO_3$.

Antiferroelectricity (without superposed ferroelectricity) is possible in all crystal classes in which there are no polar directions, that is in 22 out of the total 32 classes. These 22 classes include 12 classes without centres of symmetry.

(*b*) **The Influence of an Electric Field.** If a strong electric field is applied to a polycrystalline specimen of $PbZrO_3$ at temperatures somewhat below the Curie point a double hysteresis loop arises, similar to that observed in ferroelectrics above the Curie point. However in the former case the part of the curve between the loops corresponds to the antiferroelectric state while in the latter case it corresponds to the paraelectric. Thus in lead zirconate the electric field produces a transition from the antiferroelectric state to the ferroelectric. From the curve of the dependence of the transition temperature on field strength $T_c = T_c(E_c)$ (Fig. 140) we can determine the difference in energy between the ferro- and antiferroelectric phases. This difference is found to be small, for example at 230° C, the heat of transition in lead zirconate is $q = 4$ cal/mole. X-ray investigations in a field strong enough to retain the ferroelectric modification show that it is rhombohedral with a cell somewhat extended along the body diagonal ($\alpha < 90°$).

(*c*) **The Influence of Isomorphic Substitution.** Fig. 142 shows the results of X-ray analysis of the system $PbZrO_3$-$PbTiO_3$ at room temperature. In this system one of the components is ferro- and the other antiferroelectric. At room temperature one of the components is tetragonal and the other has a rhombic cell, so a continuous series of solid solutions is not formed. At comparatively small concentrations of the titanate there is a rhombohedral ferroelectric phase similar to the phase which is found at the ferroelectric transition in pure $PbZrO_3$ in an electric field. At large concentrations of the titanate there is a second ferroelectric phase based on a solid solution of

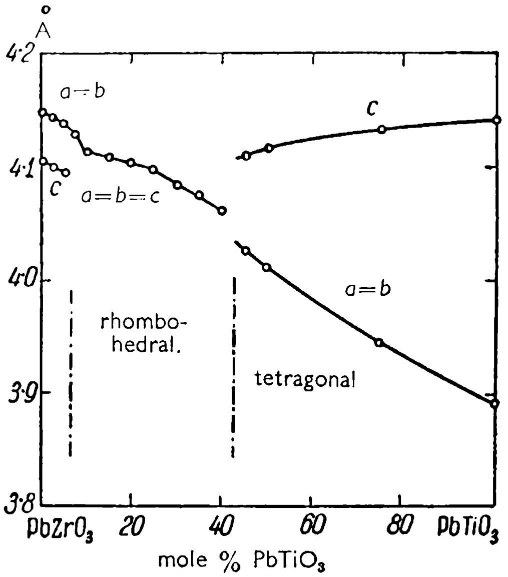


FIG. 142
Changes in the unit cell parameters as a function of concentration in the $PbZrO_3 - PbTiO_3$ system at room temperature.

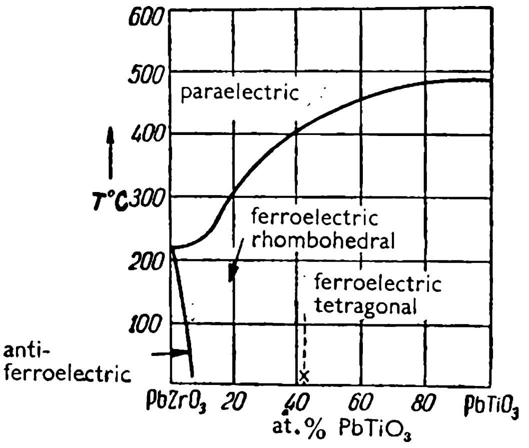


FIG. 143. Phase diagram of the $PbZrO_3 - PbTiO_3$ system.

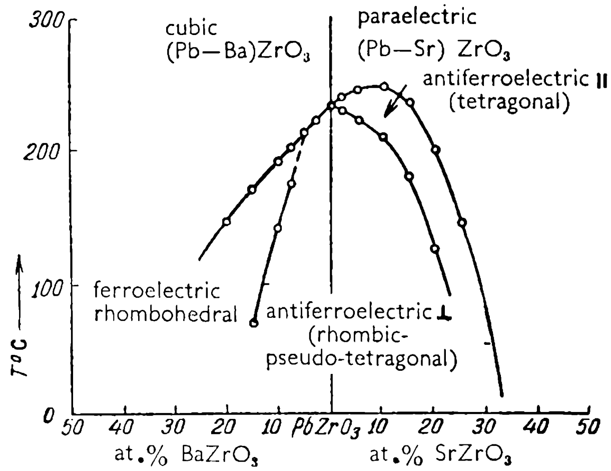


FIG. 144. Phase diagrams of the systems (Pb, Ba) ZrO_3 and (Pb, Sr) ZrO_3.

zirconate in tetragonal $PbTiO_3$. The composition diagram for $PbZrO_3$-$PbTiO_3$ is shown in Fig. 143: at sufficiently high temperatures there is a continuous series of solid solutions Pb (Ti, Zr) O_3 with cubic structure (paraelectric phase). At lower temperatures ferro- and antiferroelectric phases are found. As the concentration of titanate increases, T_c of the antiferroelectric transition falls sharply while T_c of the ferroelectric transition increases.

There is a similar succession of phases in the system (Pb, Ba) ZrO_3 (Fig. 144) while in the system (Pb, Sr) ZrO_3 we have two antiferroelectric phases: rhombohedral and tetragonal. Both these phases have been shown to exist in pure $PbHfO_3$ which has two antiferroelectric Curie points (T_c = 215° and 165° C).

(*d*) **Other Antiferroelectrics.** Antiferroelectricity has recently been discovered amongst four groups of compounds: (1) perovskite, (2) potassium phosphate, (3) silver periodate, (4) the alums (Table 56). In sodium niobate there are two paraelectric and two antiferroelectric phases and one phase with superposed ferro- and antiferroelectric states. In tungsten trioxide, two antiferroelectric and one ferroelectric phase have been established. In the potassium phosphate group, antiferroelectric transitions have been discovered in the phosphates and arsenates of ammonia; here the dimensions of the unit cell of the crystal are not changed in this transition but the cell changes from body-centred to primitive; this is analogous to the changes of the Bravais lattice during atomic ordering in β-CuZn. The rhombic phases of the phosphate and arsenate of ammonia are examples of antiferroelectric phases without centres of symmetry; here, as in the case of potassium phosphate and the tartrates, the hydrogen atoms take part in the transition. This is shown by the fact that deuteration of the compound greatly increases the Curie point. In the alums, antiferroelectric transitions have not been extensively studied.

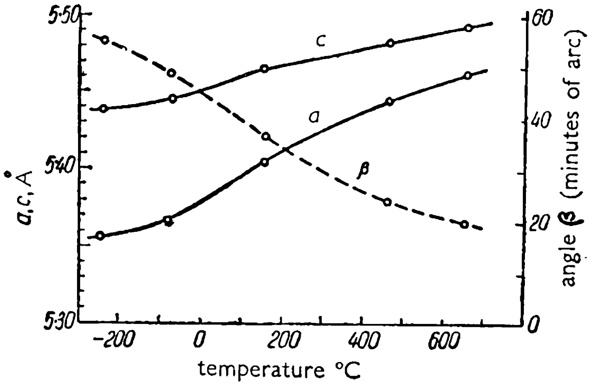


FIG. 145. Dependence on temperature of the parameters *a* and *c* of the rhombic superlattice and of the monoclinic angle β in $CaTiO_3$.

Compressions of a cubic lattice and the appearance of superlattices are known in a wide variety of other examples of the perovskite group including perovskite $CATiO_3$. $CaTiO_3$ has a superlattice ($\sqrt{2}a \times 2a \times \sqrt{2}a$). As the temperature is raised, the lattice tends to cubic (Fig. 145), but up to 900° C traces of the superlattice are retained and there is no phase transition. $NaTaO_3$ and $CdTiO_3$ have similar superlattice cells but in sodium tantalate the transition to the cubic structure takes place at T_c = 475° C, while in cadmium titanate there is no such transition. Perovskite and cadmium titanate are antiferroelectrics with high transition temperatures (antielectrets).

Generalisation of the Phenomenological Theory of Spontaneous Polarisation

(*a*) **Free Energy of a Crystal consisting of two Polarised Sub-lattices.** Even before the discovery of antiferroelectrics, Kittel (1951), starting from the analogy with antiferromagnetics, gave a generalised expression for the free energy of a dielectric which

Table 56

Crystals with Antiferroelectric Phases

Name of the crystal and formula	Temperatures of the phase transitions °C, symmetry and type of phase (P = paraelectric, A = antiferroelectric, F = ferroelectric)	Reference
Lead zirconate ($PbZrO_3$)	cubic P / 230° / rhombic A	1
Lead hafnate ($PbHfO_3$)	cubic P / 215° / tetrag. AI / 163° / rhombic AII	2
Sodium niobate ($NaNbO_3$)	cubic PI / 638° / tetrag. PII / 500° / rhombic AI / 350° / rhombic AII / 210° / monocl. AIII	3
Sodium tantalate ($NaTaO_3$)	cubic P / 475° / rhombic A	17
Calcium titanate ($CaTiO_3$) Cadmium titanate ($CdTiO_3$)	rhombic A?	16 17
Tungsten trioxide (WO_3)	tetrag. P / 900° / tetrag. AI / 700° / monocl. AII / −10° / F?	4–7
Ammonium phosphate $(NH_4)H_2PO_4$ Ammonium deuterophosphate $(ND_4)D_2PO_4$	tetrag. $J\bar{4}d2$ P / −125° / rhombic $P2_12_12_1$ A tetrag. $J\bar{4}d2$ P / −31° / rhombic $P2_12_12_1$ A	8–10
Ammonium arsenate $(NH_4)H_2AsO_4$ Ammonium deuteroarsenate $(ND_4)D_2AsO_4$	tetrag. $J\bar{4}d2$ P / −57° / rhombic $P2_12_12_1$ A tetrag. $J\bar{4}d2$ P / −31° / rhombic $P2_12_12_1$ A	11–12
Silver periodate ($Ag_2H_3IO_6$)	trigonal $R3c$ P / ? / trigonal A	13
Ammonium periodate $(NH_4)_2H_3IO_6$	trigonal $R3c$ P / ? / trigonal A	14
Alums $(R(H_2O)_6)(M(H_2O)_6)(SO_4)$	cubic P / T_c / A depends on type of R and M	15

REFERENCES TO TABLE 56

1. E. Sawaguchi, H. Maniwa and S. Hoshino: *Phys. Rev.*, **83**, 1078 (1951).
2. G. Shirane and R. Pepinsky: *Phys. Rev.*, **91**, 812 (1953).
3. P. Vousden: *Acta Cryst.*, **4**, 373 (1951).
4. H. Braekken: *Zeits f. Krist.*, **78**, 484 (1931).
5. R. Ueda, T. Tchinokawa and T. Tchinokawa: *Phys. Rev.*, **82**, 563 (1951).
6. W. L. Kehl, R. G. Hay and D. Wahl: *Phys. Rev.*, **82**, 774 (1951).
7. R. Ueda and J. Kobayashi: *Phys. Rev.*, **91**, 1565 (1953).
8. B. C. Frazer and R. Pepinsky: *Acta Cryst.*, **6**, 273 (1953).
9. H. A. Levy, S. W. Peterson and S. H. Simonsen: *Phys. Rev.*, **93**, 1120 (1954).
10. G. E. Bacon and R. S. Pease: *Proc. Roy. Soc.*, **220A**, 397 (1953); **230A**, 359 (1955).
11. E. A. Wood, W. J. Merz and B. T. Matthias: *Phys. Rev.*, **87**, 544 (1952).
12. B. C. Frazer: *Phys. Rev.*, **91**, 246 (1953).
13. H. Granicher and W. Petter: *Helv. Phys. Acta*, **27**, 216 (1954).
14. G. Busch, W. Känzig and W. M. Meier: *Helv. Phys. Acta*, **26**, 385 (1953).
15. R. Pepinsky, F. Jona and F. Shirane: *Phys. Rev.*, **102**, 1181 (1956).
16. H. F. Kay and P. C. Bailey: *Acta Cryst.*, **10**, 219 (1957).
17. H. F. Kay and J. L. Mills: *Acta Cryst.*, **10**, 213 (1957).

described the antiferroelectric state. To take the simplest case, consider a crystal the structure of which is the superposition of two identical sub-structures a and b. We may break down the polarisation of the crystal into the sum of the polarisations of the superlattices which we call P_a and P_b. The internal fields which exist on the lattice sites of these sub-structures will be

$$\begin{aligned} E_a &= E+s_1P_a-s_2P_b, \\ E_b &= E+s_1P_b-s_2P_a, \end{aligned} \qquad \ldots [9.80]$$

where E is the external field and s_1 and s_2 are structural coefficients, which are positive for an antiferroelectric state. The polarisations of the sub-lattices are proportional to the internal fields; thus

$$E_a = kP_a; \quad E_b = kP_b, \qquad \ldots [9.81]$$

where the coefficient of proportionality is related to the polarisability α of the particles of the crystal by the relation $k = (N\alpha)^{-1}$. Substituting [*81*] into formula [*80*] we obtain equations for the polarisability of the sub-lattices

$$(k-s_1)P_a+s_2P_b = E, \qquad s_2P_a+(k-s_1)P_b = E. \qquad \ldots [9.82]$$

When $E = 0$ equations [*82*] have non trivial solutions if their determinant is zero

$$\begin{vmatrix} k-s_1 & s_2 \\ s_2 & k-s_1 \end{vmatrix} = 0. \qquad \ldots [9.83]$$

From [*83*] we obtain two roots for k

$$k_1 = s_1-s_2; \quad k_2 = s_1+s_2. \qquad \ldots [9.84]$$

while [*82*] gives

$$\frac{P_a}{P_b} = -\frac{s_2}{k-s_1}. \qquad \ldots [9.85]$$

Now substituting the values of k from [*84*] we find for k_1, $P_a = P_b$ and for k_2, $P_a = -P_b$.

The former solution describes a ferroelectric state and the second an antiferroelectric state. Let us now find the change of electric susceptibility μ at the transition point. Adding equations [*82*] we obtain

$$(k-s_1+s_2)(P_a+P_b) = 2E, \qquad \ldots [9.86]$$

from which

$$\mu = \frac{P_a+P_b}{E} = \frac{2}{k-s_1+s_2}. \qquad \ldots [9.87]$$

At the ferroelectric Curie point, the susceptibility becomes infinitely large: $(\mu)_{k_1} \to \infty$ ('the polarisation catastrophe'), while at the antiferroelectric Curie point the susceptibility still has a finite value $(\mu)_{k_2} = \frac{1}{s_2}$.

We can write the free energy of a crystal which consists of two polarised sublattices in the form

$$F(P_a, P_b) = \alpha_1(P_a^2+P_b^2)+\alpha_2 P_a P_b+\beta(P_a^4+P_b^4)+\gamma(P_a^6+P_b^6)+\ldots \qquad \ldots [9.88]$$

whence we obtain the field strength

$$\frac{\partial F}{\partial P_a} = E = 2\alpha_1 P_a+\alpha_2 P_b+4\beta P_a^3+6\gamma P_a^5+\ldots \qquad \ldots [9.89]$$

(*b*) **Antiferroelectric Second Order Transition.** If we cut off the series in (88) and [*89*] at their third terms we obtain for the spontaneous polarisation of the substructures ($P_a = -P_b$) in the absence of an external field ($E = 0$)

$$P_{as}^2 = P_{bs}^2 = \frac{\alpha_2-2\alpha_1}{4\beta}. \qquad \ldots [9.90]$$

If $\alpha_1 < 0$, the spontaneous polarisation changes continuously through the transition point from the para- to the ferroelectric state. At the Curie point

$$P_{as}^2 = 0, \quad \alpha_2 = 2\alpha_1. \qquad \ldots [9.91]$$

We now arrange that a weak external field ΔE acts on the crystal producing a change in polarisation

$$\Delta P = \Delta P_a+\Delta P_b.$$

Differentiating [*89*] and the analogous expression for $E = \frac{\partial F}{\partial P_b}$, and adding the results of differentiation we obtain

$$2\Delta E = (2\alpha_1+\alpha_2+12\beta P_{as}^2)\Delta P, \qquad \ldots [9.92]$$

whence

$$\mu = \frac{\Delta P}{\Delta E} = \frac{2}{2\alpha_1+\alpha_2+12\beta P_{as}^2}. \qquad \ldots [9.93]$$

Below the Curie point, equation [*90*] is valid. Substituting the value [*93*] we obtain

$$\mu^{-} = \frac{1}{2(\alpha_2-\alpha_1)}, \qquad \ldots [9.94]$$

which at the Curie point and with $\alpha_2 = 2\alpha_1$ gives

$$\mu_{T_c}^{-} = \frac{1}{\alpha_2}. \qquad \ldots [9.95]$$

Above the Curie point $P_{as}^2 = 0$ and from [*93*] we have

$$\mu^{+} = \frac{2}{2\alpha_1+\alpha_2}. \qquad \ldots [9.96]$$

At the Curie point we obtain

$$\mu_{T_c}^{+} = \frac{1}{\alpha_2}, \qquad \ldots [9.97]$$

which agrees with [*95*]. Thus at the Curie point the susceptibility and dielectric constant are continuous.

Now assume that the coefficient of the first term of the series in [*89*] depends linearly on temperature

$$\alpha_1 = \tfrac{1}{2}\alpha_2 + \lambda(T-T_c), \qquad \ldots [9.98]$$

where $\lambda = \dfrac{1}{C}$ is the reciprocal of the Curie constant. From [*91*] and [*98*] we obtain a parabolic dependence of the spontaneous polarisation of the sub-lattices on temperature

$$P_{as}^2 = P_{bs}^2 = \frac{\lambda}{2\beta}(T_c-T). \qquad \ldots [9.99]$$

For the susceptibility above and below the transition temperature we have

$$\mu = \frac{1}{\alpha_2+\lambda(T-T_c)}, \ (T>T_c); \quad \mu = \frac{1}{\alpha_2-2\lambda(T_c-T)}, \ (T<T_c). \qquad \ldots [9.100]$$

If $\lambda > 0$, the antiferroelectric region lies below T_c and the curve of $\varepsilon(T)$ passes through a maximum. If $\lambda < 0$ the antiferroelectric region lies above T_c and the curve of $\varepsilon(T)$ passes through a minimum at T_c.

(*c*) **Antiferroelectric First Order Transition.** In the expression for the free energy include the term of sixth order. The spontaneous polarisation ($E = 0$) in the antiferroelectric region ($P_a = -P_b$) in this case is determined by the equation

$$6\gamma P_{as}^4 + \beta P_{as}^2 + (2\alpha_1 - \alpha_2) = 0. \qquad \ldots [9.101]$$

We obtain a second equation from the condition $F = 0$ which gives

$$2\gamma P_{as}^4 + 2\beta P_{as}^2(2\alpha_1 - \alpha_2) = 0. \qquad \ldots [9.102]$$

Equations [*101*] and [*102*] lead to the solutions (at $T = T_c$)

$$P_{as}^2 = \frac{\alpha_2 - 2\alpha_1}{\beta}, \quad P_{as}^4 = \frac{2\alpha_1 - \alpha_2}{2\gamma}. \qquad \ldots [9.103]$$

Hence we obtain the relation between the coefficients of the series at the Curie point

$$\beta^2 = 2\gamma(2\alpha_1 - \alpha_2). \qquad \ldots [9.104]$$

We also obtain an expression analogous to [*93*] for the susceptibility

$$\eta = \frac{\Delta P}{\Delta E} = \frac{2}{2\alpha_1 + \alpha_2 + 12\beta P_{as}^2 + 30\gamma P_{as}^4}. \qquad \ldots [9.105]$$

Below the Curie point equation [*103*] is valid, thus

$$\eta^- = \frac{1}{4\alpha_1 - \alpha_2}. \qquad \ldots [9.106]$$

Above the Curie point $P_{as}^2 = 0$, so

$$\eta^+ = \frac{2}{2\alpha_1 + \alpha_2}. \qquad \ldots [9.107]$$

As we would expect from the shape of the curve of free energy, the transition from the para- to the antiferroelectric state in this case is discontinuous since the condition for a continuous transition [*91*] is not fulfilled. The susceptibility and dielectric constant are thus discontinuous at the Curie point.

The Simplest Model of a Ferro- and Antiferroelectric

It was shown above (§2) that for a simple cubic lattice with freely rotating dipoles the antipolar structure (Fig. 126*a*) was energetically more favourable than the polar structure. Thus for such a dipole system antiferroelectricity is more probable than ferroelectricity. Takagi considered a model based on a b.c.c. structure, on the corners of which freely rotating dipoles were distributed each having a moment p, while on the centres there were ions (Fig. 146). The dipoles and ions can be mutually polarised and the respective polarisabilities are then denoted by α_1 and α_2, which must not be confused with the coefficients in the series for the free energy. We divide the dipole lattice into two sub-lattices. The polarisation of the crystal can be thought of as the sum of three parts: $P = P_1 + P_2 + P_3$, where P_1 and P_2 are the polarisation of the dipolar sub-lattices

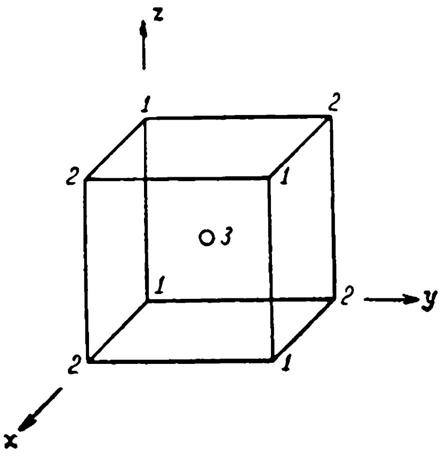


FIG. 146

Model of a cubic dielectric. Dipoles (1, 2) are situated on the corners of the cell and polarisable ions in the centre of the cell (3).

and P_3 the polarisation of the ionic sub-lattice. Analysis of the free energy of the system shows that in the absence of an external field, two solutions are possible with spontaneous polarisations of the sub-lattices differing from zero. One of them corresponds to an antiparallel arrangement of dipoles and the other to a parallel arrangement. For the antiferroelectric case, the Curie temperature and polarisation of the sub-lattice are determined by the formulae:

$$T_a = \frac{p_0^2}{6kv} \cdot \frac{s_{11}-s_{12}}{\left[1-\frac{\alpha_1}{2v}(s_{11}-s_{12})\right]^2},$$

$$P_1 = -P_2 = \frac{p_0 L(x)}{2v} \times \frac{1}{1-\frac{\alpha_1}{2v}(s_{11}-s_{12})}; \quad P_3 = 0, \qquad \ldots [9.108]$$

where v is the volume of the cubic unit cell, $L(x)$ is the Langevin function, and s_{11} and s_{12} are structural coefficients

$$(s_{11} = s_{22} = \tfrac{1}{3}+0{\cdot}426; \quad s_{12} = \tfrac{1}{3}-0{\cdot}426; \quad s_{13} = s_{23} = s_{33} = \tfrac{1}{3}).$$

For the ferroelectric case

$$T_s = \frac{p_0^2}{6kv} \cdot \frac{2}{3\left[1-\frac{(\alpha_1+\alpha_2)}{3v}\right]};$$

$$P_1 = P_2 = \frac{p_0 L(x)}{2v} \, \frac{1-\frac{\alpha_2}{3v}}{1-\frac{\alpha_1+\alpha_2}{3v}}; \quad P_3 = \frac{p_0 L(x)}{2v} \, \frac{\frac{2}{3}\frac{\alpha_2}{v}}{1-\frac{\alpha_1+\alpha_2}{3v}}. \qquad \ldots [9.109]$$

The relative stability of both configurations is determined by the parameter $Q = \frac{T_s}{T_a}$. when $Q<1$ the antiferroelectric configuration is stable and when $Q>1$ the ferroelectric configuration is stable. The parameter Q depends only on the polarisability of the components of the model

$$Q = \frac{T_s}{T_a} = 1{\cdot}76\left(\frac{2v-0{\cdot}85\,\alpha_1}{3v-\alpha_1-\alpha_2}\right)^2. \qquad \ldots [9.110]$$

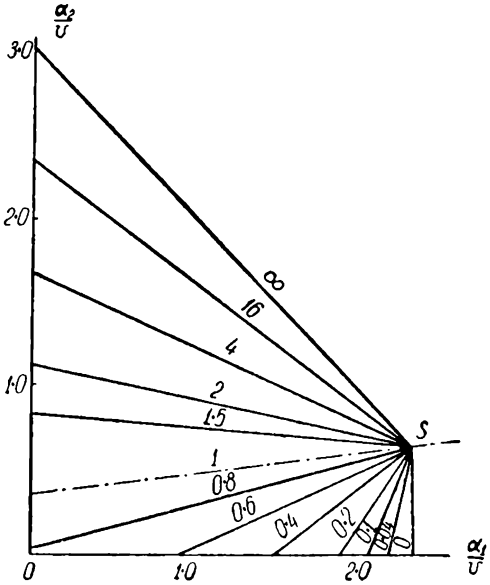


FIG. 147

Curves of constant values of $Q = \frac{T_s}{T}$; in the region lying above the dotted line $Q = 1$, the ferroelectric phase is more stable. In the region lying below, the antiferroelectric phase is more stable.

For a given value of Q, equation [*110*] is represented by a straight line on the $(\alpha_1\ \alpha_2)$ plane. This is shown in Fig. 147 for various values of Q. Below $Q = 1$ the antiferroelectric solution is stable, and above it the ferroelectric solution is stable. In this model the ionic polarisability is the important factor in establishing the ferroelectric state. If polarisation is increased the stability of the ferroelectric state also increases. These results are the same also for a model in which the rotating dipoles are replaced by induced dipoles.

The Structural Mechanism of Dielectric Phase Transitions

A knowledge of the atomic structure of ferro- and antiferroelectrics above and below the transition points is necessary for an explanation of the structural mechanism of the transition and the development of a microscopic theory of the phenomena. Analysis of the phases of ferro- and antiferroelectrics, the determination of their symmetry and the measurements of the deformations of the cells as a result of dielectric phase transitions are all carried out by means of X-rays. To explain the patterns of atomic displacements and determine their magnitude (accurate localisation of light atoms in the presence of heavy metal atoms) requires the application of neutron diffraction.

(*a*) **Structure of the Tetragonal Modifications of $BaTiO_3$ and $PbTiO_3$.** The cell of the tetragonal modification of barium titanate is $a_t = 3{\cdot}980$; $c_t = 4{\cdot}026$ Å; $c_t/a_t = 1{\cdot}011$. If we compare this with the dimensions of the cubic cell, extrapolated to room temperature, this corresponds to extension along the *c*-axis by 0·6% and compression along the *a*-axis by 0·25%. Thus there is a small increase in volume at the ferroelectric transition. Deformation of the cell is accompanied by small displacements of the ions from the symmetrical positions which they occupy in the cubic modification. In $PbTiO_3$ the deformation of the cell is larger ($a_t = 3{\cdot}90$; $c_t = 4{\cdot}15$ Å; $c_t/a_t = 1{\cdot}060$). The structure of a tetragonal modification of a perovskite ABO_3 type is determined by the following positions of the atoms:

$$\mathrm{A}\,(0, 0, 0+\delta Z_A),\quad \mathrm{B}\,(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}+\delta Z_B)$$
$$\mathrm{O}_{\mathrm{I}}\,(\tfrac{1}{2}, \tfrac{1}{2}, \delta Z_{\mathrm{O_I}}),\quad 2\mathrm{O}_{\mathrm{II}}\,(\tfrac{1}{2}, 0, \tfrac{1}{2}+\delta Z_{\mathrm{O_{II}}};\ \ 0, \tfrac{1}{2}, \tfrac{1}{2}+\delta Z_{\mathrm{O_{II}}}),$$

where δZ_i is the displacement of the ions along the axis of Z. We now introduce the parameter

$$t = \frac{r_A + r_B}{\sqrt{2}\,(r_B + r_0)}, \qquad \ldots [9.111]$$

which determines the contact of ions of types A and B with ions of oxygen. Note that $t = 1$ corresponds to close contact of ions while $t > 1$ and $t < 1$ correspond to free movement of cations of types B and A respectively. Using this parameter, we can connect the origin of coordinates with the cations which are in close contact with the anion carcass of the crystal. In $BaTiO_3$ $t > 1$ and for $PbTiO_3$ $t < 1$. Therefore for these compounds we have to use somewhat different origins, but for the sake of simplicity in considering ionic displacements in both crystals we can refer the origin to the oxygen ions which lie on the side faces: $\delta Z_{\mathrm{O_{II}}} = 0$. X-ray and neutron diffraction

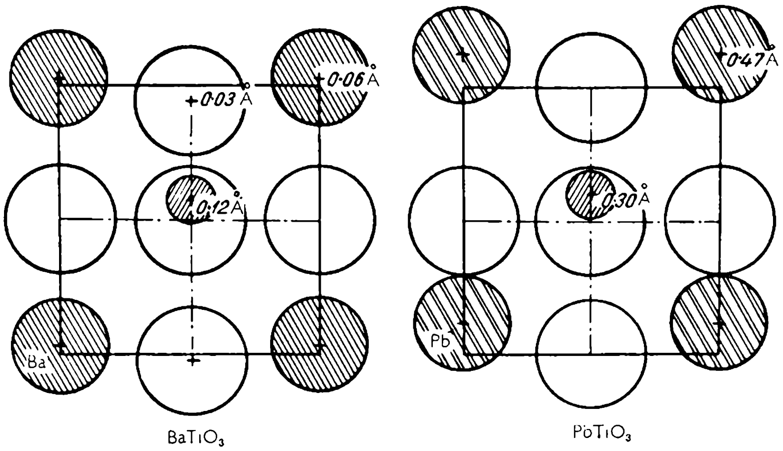


FIG. 148. Displacements of ions in the tetragonal modifications of $BaTiO_3$ and $PbTiO_3$.

results from single domain crystals in a constant electric field give the following data for ionic displacements:

ABO_3	δZ_A	δZ_B	δZ_{O_I}	$\delta Z_{O_{II}}$
$BaTiO_3$	0·06	0·12	−0·03	(0·00) Å
$PbTiO_3$	0·47	0·30	0·00	(0·00)

The atomic displacements in both crystals are shown schematically in Fig. 148. The O_I ions are not much displaced and this shows itself in the comparatively small deformation of the oxygen lattice. Thus the oxygen octahedra retain their regular shape in the tetragonal modification of barium titanate. The displacements of the cations in $PbTiO_3$ are 5-6 times greater than in $BaTiO_3$ and this corresponds to the difference in the degree of tetragonal deformation (6 and 1%). Again the pattern of cation displacement is different in the two compounds: in $PbTiO_3$ cations of type *A* are most displaced, while in $BaTiO_3$ the cations of type *B* are most displaced. This agrees with the difference in the closeness of cation-oxygen contact in the cubic modifications, that is with what we measure by the factor *t*.

Table 57

Interatomic Spacings M-O in $BaTiO_3$ and $PbTiO_3$ at Room Temperature (Å)

Cation		Coordination Polyhedron	Anion	$BaTiO_3$ $t>1$	$PbTiO_3$ $t<1$
A	Ba or Pb	Icosahedron	O_{II} upper (4 Bonds)	2·80	2·53
			O_I (4 Bonds)	2·82	2·80
			O_{II} lower (4 Bonds)	2·88	3·20
B	Ti	Octahedron	O_I upper	1·86	1·78
			O_{II} (4 Bonds)	2·00	1·98
			O_I lower	2·17	2·38
Lattice Constants			*a*	3·99	3·90
			c	4·03	4·15
Parameter			*t*	1·03	0·98

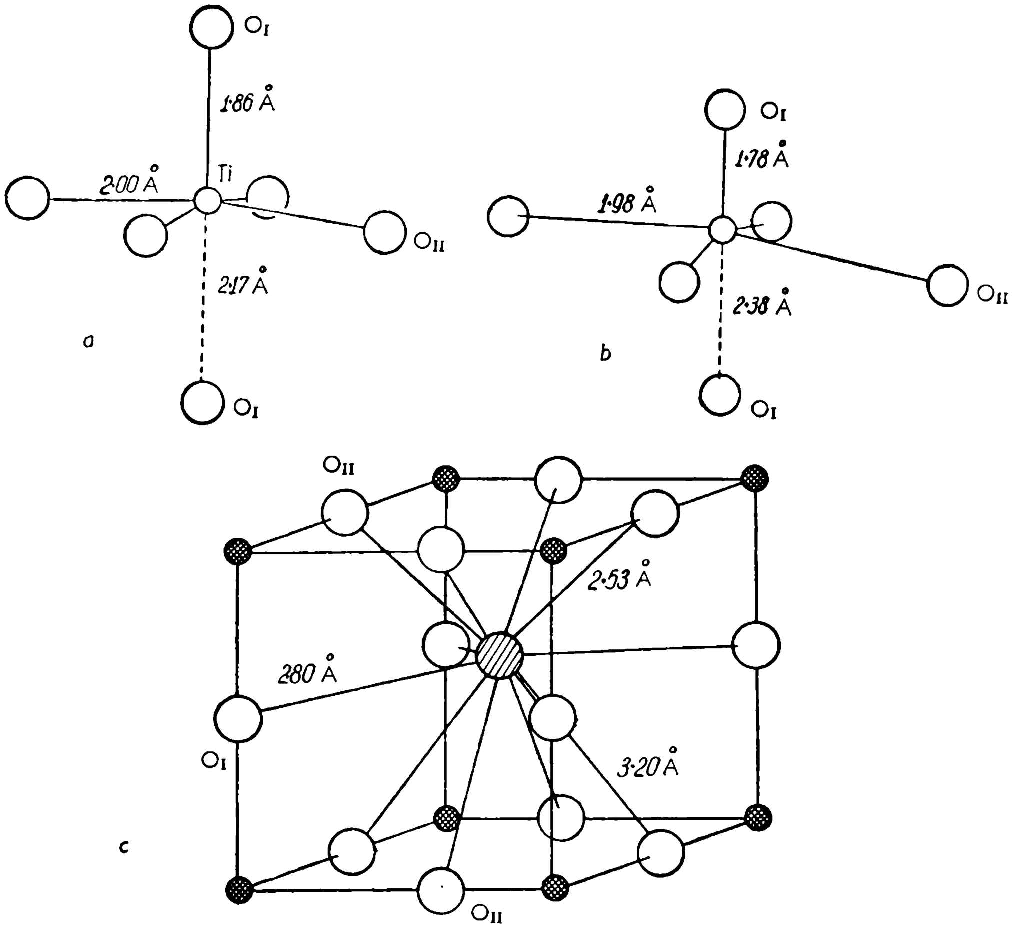


FIG. 149. Structural configurations for the A and B cations in tetragonal modifications of $BaTiO_3$ and $PbTiO_3$ for (*a*) Ti in $BaTiO_3$, (*b*) Ti in $PbTiO_3$ and (*c*) Pb in $PbTiO_3$.

Because of deformation, the displacements of ions δZ_i reckoned at the phase transitions, do not describe all the changes of the structural configuration for the ions of the various types. The actual structural configurations are determined by the interatomic spacings and by the valence angles. Values of the interatomic spacings M-O for both compounds at room temperature are given in Table 57.

In the cubic modification, ions of type *A* are situated in the centres of regular icosahedra and ions of type *B* in the centres of regular octahedra. In the tetragonal modifications the cubic symmetry of these polyhedra breaks down for, as a result of the movement of the cations along one of the cube axes in the octahedra, one of the M-O spacings is greatly reduced and the other correspondingly increased (Fig. 149*a* and *b*). In the icosahedra four M-O spacings are reduced and four others are increased (Fig. 149*c*). The reductions in the spacings are of the order of tenths of an Angstrom which is not negligible compared with the sums of the ionic radii (Ti-O 1·96 Å; Pb-O 2·78 Å). Movements of the cations lead to a change in the coordination numbers for the octahedra from 6 to $1+4+1$ and for the icosahedra from 12 to $4+4+4$. Lowering of the coordination number, without changing the type of chemical binding, is accompanied by a reduction in interatomic spacing (Table 37, pages 192-3). This effect is also evident in ferroelectric transitions. Displacement of the cations from the symmetric positions is accompanied by corresponding deformation of the valence angles.

Knowing the structure and effective charges of the ions we can calculate that part of the spontaneous polarisation which is due to ionic movements

$$P_i = \frac{1}{v}\sum e_i^* \delta Z_i,$$

where v is the volume of the unit cell and the summation is over all ions in the cell. If we deliberately exaggerate the calculation by including the full ionic charge (A^{+2}, B^{+4}, O^{-2}) we get

$$PbTiO_3 \quad P_i = 54\ \mu C/cm^2 \quad \text{and for} \quad BaTiO_3 \quad P_i = 17\ \mu C/cm^2.$$

The experimental value of the spontaneous polarisation is known only for barium titanate and is $P_s = 26\ \mu C/cm^2$. This is larger than the calculated value P_i, showing that there is a large contribution from the electronic polarisation. A more careful calculation of the internal field and spontaneous polarisation of $BaTiO_3$ taking

Table 58

Classification according to Crystal Chemistry of the Perovskite Compounds which have Dielectric Transitions

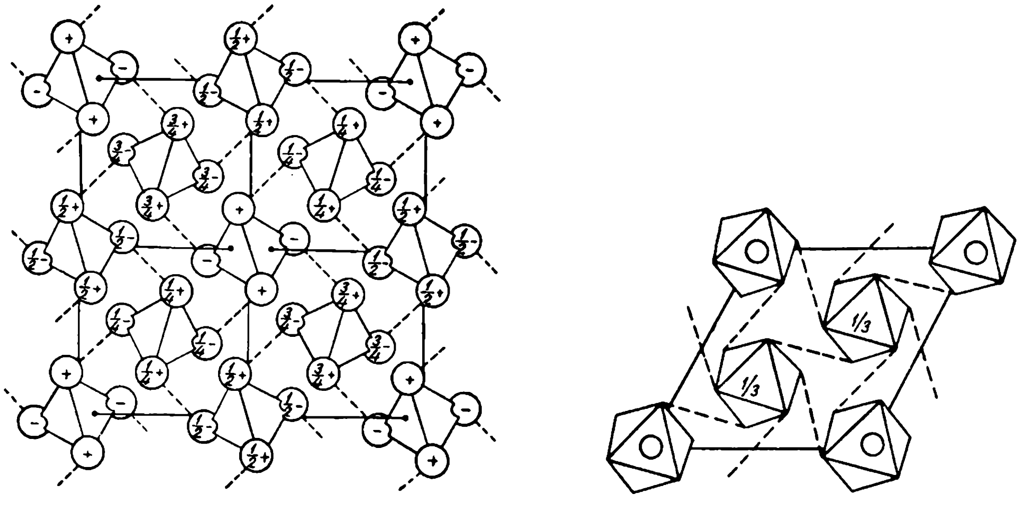
Crystal ABO_3	t	Largest displaced cation	Phases and transition points in °C (P, F, A denote respectively para-, ferro, and antiferroelectrics)	Type of spontaneous polarisation
$BaTiO_3$	$1{\cdot}0_3$	B	cubic P/118°/tetrag F/ 0°/rhombic F/−90°/rhombohedral F	Ferro-electric
$KNbO_3$	$1{\cdot}0_5$		cubic P/435°/tetrag F/210°/ rhombic F/ −40°/rhombohedral F	
$KTaO_3$	$1{\cdot}0_0$	A	cubic P/ −260°/ ? F	
$PbTiO_3$	$0{\cdot}9_8$		cubic P/500°/tetrag F	
$SrTiO_3$†	$0{\cdot}9_6$			
$PbHfO_3$	$0{\cdot}9_2$	A	cubic P/215°/tetrag AI/ 163°/rhombic AII	Anti-ferro-electric
$PbZrO_3$	$0{\cdot}9_0$		cubic P/230°/rhombic A	
$CaTiO_3$	$0{\cdot}8_9$		>900°/rhombic A	
$CdTiO_3$	$0{\cdot}8_7$		rhomb?/ −210°/rhombic	
$NaNbO_3$	$0{\cdot}8_6$		cubic P/638°/tetrag P'/ 500°/rhombic AI/350°/ rhombic AII/ −210°/AIII	
$NaTaO_3$	$0{\cdot}8_6$		cubic P/ 475°/rhombic A	
WO_3	$0{\cdot}2_3$		tetrag P/900°/tetrag AI/ 700°/monocl. AII/ −10°/F ?	

† An induced ferroelectric in an electric field.

account of the electronic polarisability of the ions for half ionic charges gives $P_s = 29\ \mu\text{C/cm}^2$ (Venevtsev, Zhdanov *et al.*). The ionic part turns out to account for 30% of the total polarisation and the electronic for 70%.

Various theories have been put forward for barium titanate. These include: the theory of local minima, the theory of anharmonic oscillators, a quantum-mechanical theory which takes account of electron interaction in the MO_6 octahedra, and Slater's theory which is based on a model derived empirically from structural investigations. The weak points in the majority of these theories are discussed in the literature. Table 58 gives U. N. Venevtsev's results on the crystal-chemical classification of the perovskite compounds for which a dielectric transition has been established: the ferroelectrics are grouped around a value of t close to unity, while antiferroelectric properties are characteristic of compounds for which $t < 1$. The hetero-isomorphic compounds $KNbO_3$ and $BaTiO_3$ for which the value of t is somewhat greater than unity have completely analogous ferroelectric transitions. Although they are alike in valency, $BaTiO_3$ and $PbTiO_3$ differ in their ferroelectric transitions, the former having only one ferroelectric phase. For these compounds we have respectively $t > 1$ and $t < 1$, and this leads to a difference in the mechanism of transition and a difference in the stability of the individual ferroelectric phases.

(*b*) **The Structure of the Phosphates and the Periodates.** The structures of the phosphates, the arsenates which are isomorphic with them, and also the periodates are based on a single principle. The general formula of the phosphates is $M.H_3[PO_4]$ where M is an alkali metal or ammonia; the formula of the periodates is $M_2.H_3[IO_6]$ where M is Ag or ammonia; in both cases the basic structure consists of a three dimensional packing of complex anions: tetrahedra of PO_4^{-3} (Fig. 150*a*) or octahedra of IO_6^{-5} (Fig. 150*b*) each bound to the other by hydrogen bonds. The number of oxygen atoms in each structure is less by a factor two than the number of vertices of the corresponding polyhedra. Thus in the phosphates it is two per polyhedron and



a *b*

FIG. 150. Packing of complex hydrogen bonded anions: (*a*) tetragonal packing of the tetrahedra PO_4^{-3} in the phosphates MH_2PO_4, (*b*) rhombohedral packing of the octahedra IO_6^{-5} in the periodates $M_2H_3IO_6$.

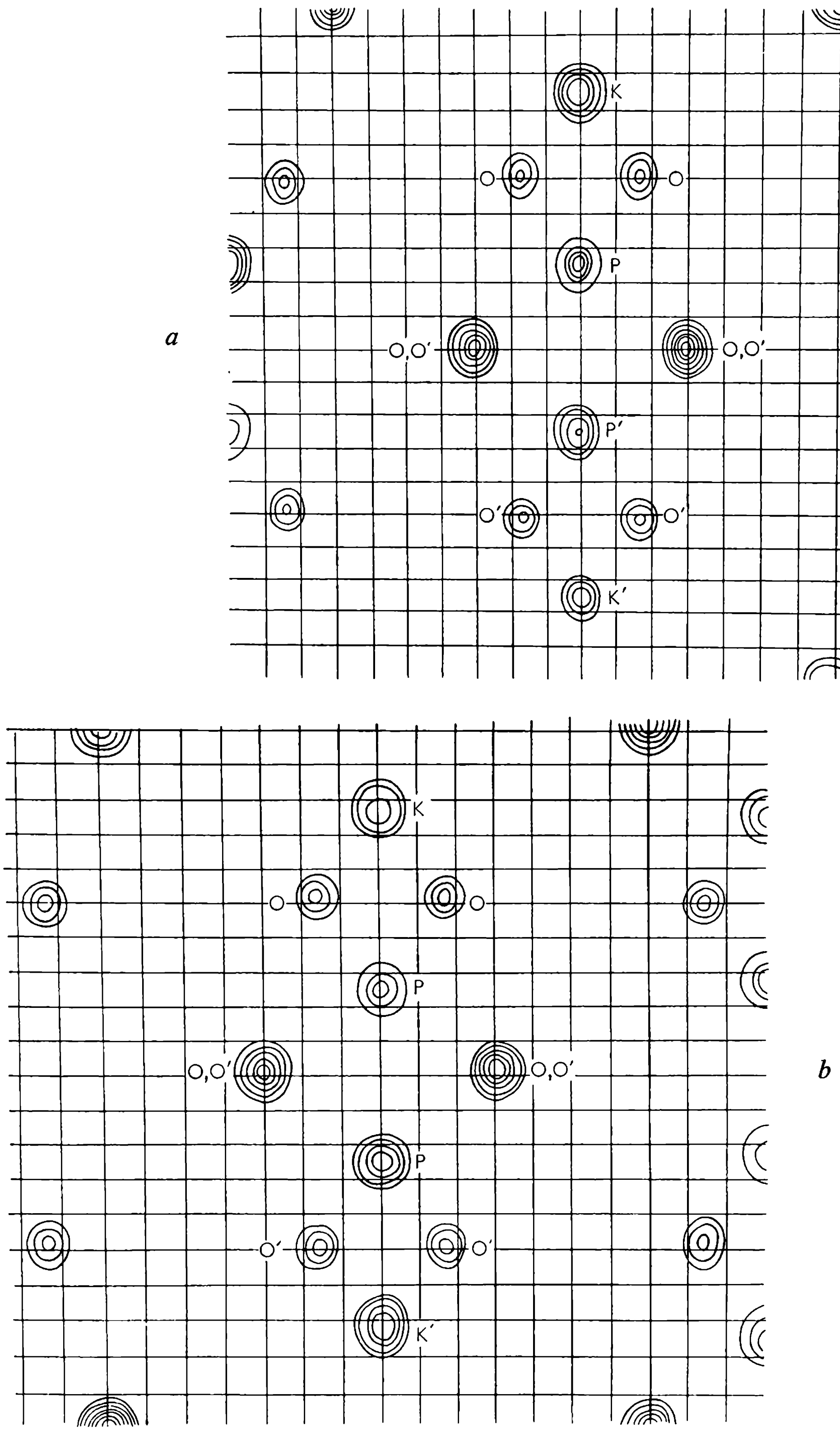


FIG. 151. Projections of the electron density on the (100) face of a KH_2PO_4 crystal, (*a*) above (126° K) and (*b*) below (116° K) the Curie point. Both from the X-ray results of Frazer and Pepinsky.

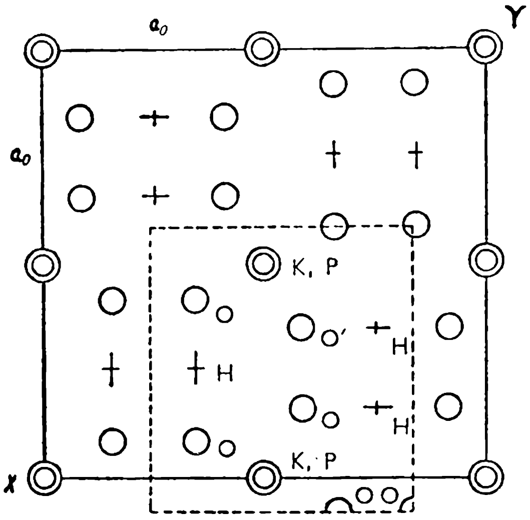
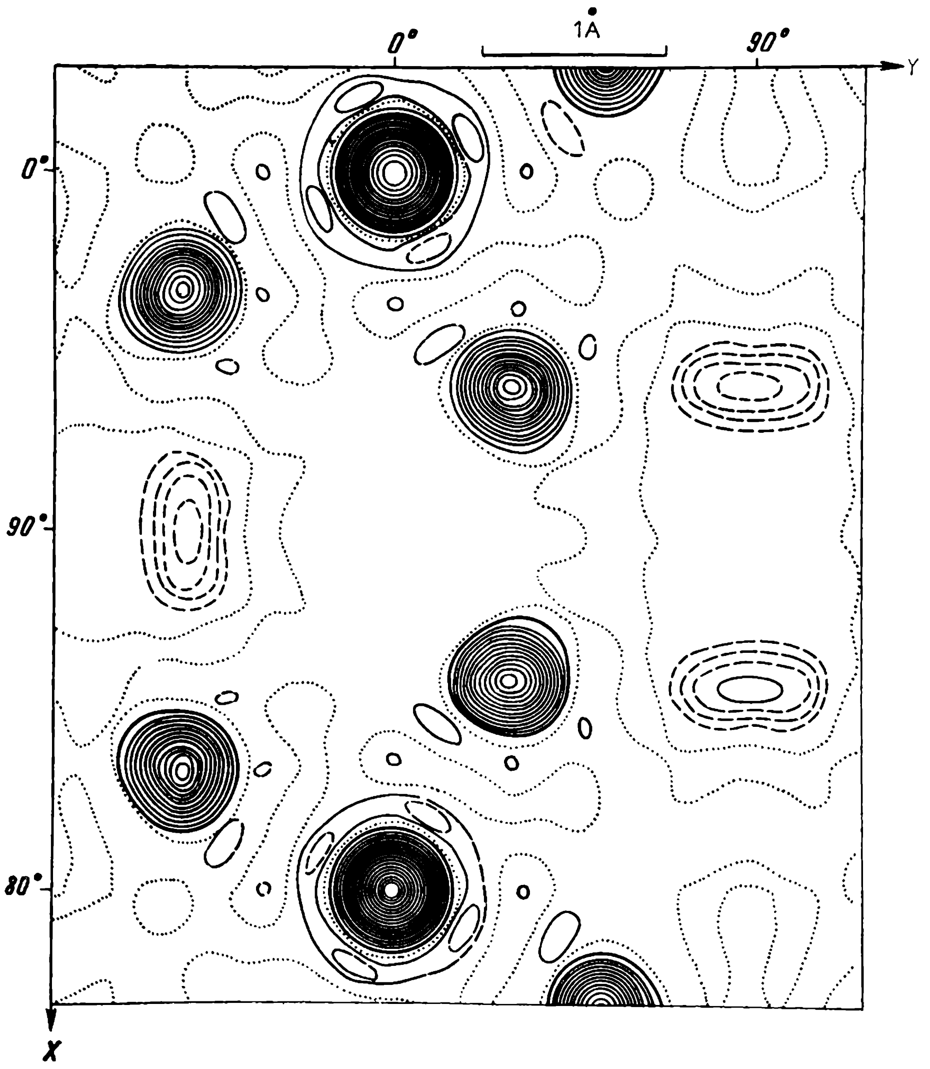


FIG. 152

a



b

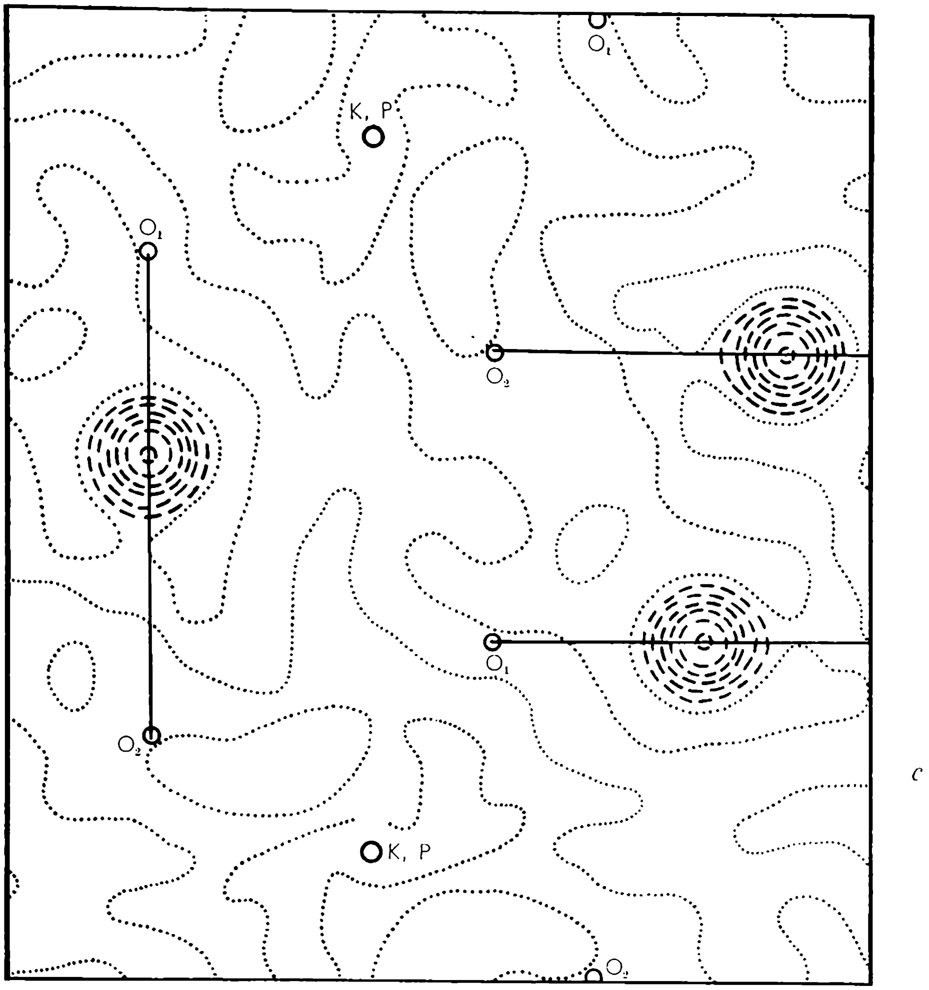


FIG. 152. Projections on the (001) face of (*a*) the structure of $K_2H_4PO_4$ (the dotted region is that shown in *b* and *c*); (*b*) the density of neutron scattering at room temperature (all nuclei are shown); (*c*) the density of neutron scattering at − 180° C (only protons are shown).

in the periodates three. The surplus negative valency is compensated by the corresponding number of cations which are distributed within the cation framework.

(*c*) **Structural Changes in KH_2PO_4 during the Ferroelectric Transition.** In Slater's theory the ferroelectric transition in KH_4PO_4 is related to the transition from the disordered to the ordered distribution of hydrogen atoms in hydrogen bonds. On this model, it is still not clear why a dipole moment should arise along the *c*-axis, since the oxygen bonds are directed almost perpendicular to the tetragonal axis (their inclination does not exceed 0·5°). Later structural investigations have confirmed that protons make the transition to the ordered state below the Curie point and that this is connected with horizontal displacement of the protons. Simultaneously, vertical displacements of all the other atoms take place and this leads to the appearance of a moment directed along the tetragonal axis. Fig. 151*a* and *b* show projections of the electron density on the side face of the cell above and below the Curie point; each projection shows two PO_4 tetrahedra, above and below which the potassium atoms are located. These PO_4 tetrahedra are almost regular, the potassium atoms being surrounded by eight oxygen atoms, but the spacing between two upper and two lower atoms of oxygen is greater than between edge atoms. The potassium atoms support thermal oscillations primarily along the *c*-axis and this leads to an ellipsoidal extension along the *c*-axis of the electron cloud of the potassium atoms. Similar anisotropy arises for the phosphorus atoms, and the predicted anisotropy is in agreement with the dielectric properties of the paraelectric phase. At 126° K, that is about 3° above T_c, ε_c is very large ($\sim 10^4$) while ε_a remains at a normal value. Below the Curie point, the shape of the oxygen tetrahedron is unchanged while the phosphorus atom is displaced from the centre of the tetrahedron. The upper phosphorus atom is displaced downwards by 0·03 Å while the upper potassium atom is displaced upwards by 0·05 Å. These displacements are such that the phosphorus and potassium atoms are taken further away from the two upper oxygen atoms on the vertices of the tetrahedron while the hydrogen atoms are brought nearer to these oxygens. This hydrogen ordering leads to a small extension of the hydrogen bonds O–H . . . O from 2·48 Å above the Curie point to 2·50 Å a little below T_c. The spontaneous polarisation calculated from structural data is

$$P_i = 4{\cdot}1\ \mu\text{C/cm}^2,$$

which is in agreement with the measured value $P_s = 4{\cdot}5\ \mu\text{C/cm}^2$.

Neutron diffraction data confirm these movements of the potassium and phosphorus atoms below the Curie point and show clearly the proton ordering. Fig. 152*a* shows the projection of the structure of KH_2PO_4 on the (001) face, hydrogen atoms being shown by crosses. The dotted rectangle is shown again on projections obtained by neutron diffraction at room temperature (Fig. 152*b*) and at a temperature 40° below the Curie point (Fig. 152*c*). In the latter projections, which were obtained by the difference method, hollows are seen which correspond only to protons since the scattering amplitude of neutrons by protons is negative. The negative regions in both projections are shown dotted. At room temperature, the centres of the hollows coincide with the mid-points of lines which link the oxygen atoms in neighbouring tetrahedra. These hollows have an elongated shape, being extended along the lines of the hydrogen bonds, and this can be explained in two ways: anisotropy of thermal vibrations of hydrogen atoms or random movements of hydrogen atoms in two potential wells situated on the line of each hydrogen bond. The character of the combination spectra of potassium phosphate is not changed at the Curie point, and this supports the random distribution of hydrogen atoms. Below the Curie point, as can be seen from the projection in Fig. 152*c*, there is marked ordering in the positions of the protons which are situated about 0·21 Å from the centres of the O–H . . . O bonds. Thus the distance between a hydrogen and the nearest atom of oxygen is 1·04 Å, while the distance from the further atom is 1·46 Å. This proton ordering occurs within a single domain, and to keep the crystal as a single domain the investigation is carried out in a constant field directed along the *c*-axis. If the direction of the field is reversed the protons transfer to the neighbouring positions of equilibrium and this is accompanied by a reversal in the direction of spontaneous polarisation.

(*d*) **Structural Changes in $(NH_4)H_2PO_4$ during an Antiferroelectric Transition.** Although the paraelectric modifications KH_2PO_4 and $(NH_4)H_2PO_4$ are isomorphic, the dielectric transition in these crystals is different. In ammonium phosphate, the largest changes occur in the dielectric constant ε_a, that is they take place in a direction perpendicular to the *c*-axis. This can be explained, while still retaining the general

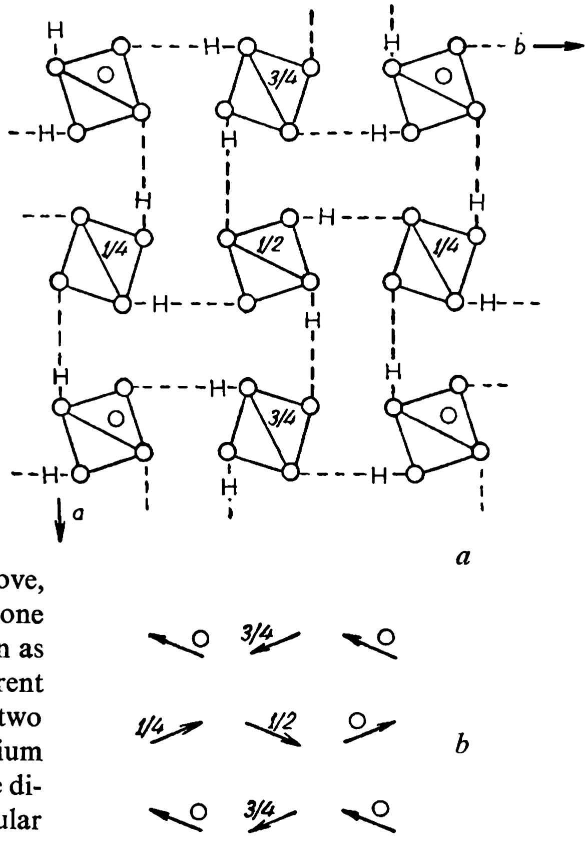


FIG. 153. (*a*) Proton ordering in $(NH_4)H_2PO_4$ below the Curie point, (*b*) directions of the elementary dipole moments.

idea of proton ordering described above, if there are two protons 'one above, one below' near to each PO_4 tetrahedron as shown in Fig. 153*a*. This is different from the scheme 'two above' or 'two below' which worked for potassium phosphate, for in the new scheme the dipole moments are directed perpendicular to the *c*-axis. The elementary dipoles form polar chains (Fig. 153*b*) strung out along one of the coordinate axes, either *X* or *Y*. In addition, neighbouring chains are directed antiparallel which corresponds to an antiferroelectric state. This scheme for the structural transition in ammonium phosphate has been confirmed by X-ray diffraction and by dielectric and optical investigations. Neutron diffraction is made difficult by the fact that large crystals of ammonium phosphate disintegrate below the Curie point.

(*e*) **The Structure of Rochelle Salt.** Rochelle salt is a salt of tartaric acid $H_2C_4H_6O_6 = (HC \,.\, OH \,.\, COOH)_2$. The acid molecule is a dimer and its structure is shown in Fig. 154. The carbon and oxygen atoms in the monomeric groups in the

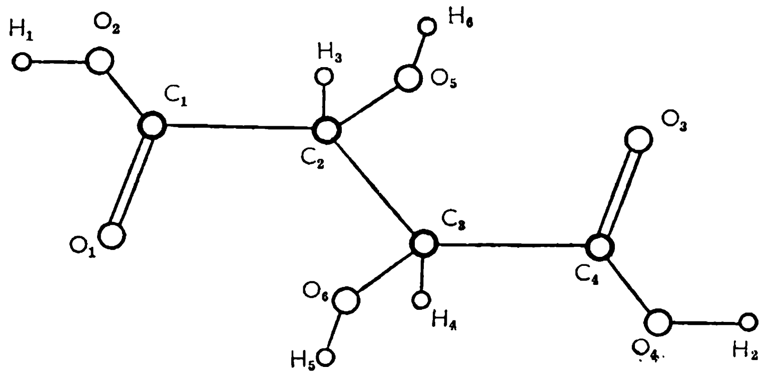


FIG. 154. Structure of the molecule of tartaric acid $H_2C_4H_4O_6$ and of its ion $C_4H_4O_6^{-2}$.

molecule of tartaric acid are rotated relative to each other by 63°. The basic structure of the salt of tartaric acid is formed by the tartrate ions $C_4H_4O_6^{-2}$ the structure of which is similar to the structure of the acid molecule if the hydrogens H_1 and H_2 are replaced by electrons (Fig. 154). The sodium potassium salt of tartaric acid $NaKC_4H_4O\,.\,4H_2O$ crystallises in two modifications: triclinic, containing two acid molecules in the unit cell and having a centre of symmetry; and rhombic, containing four acid molecules and having no centre of symmetry. The crystal structures of both modifications are very similar but the dielectric transition takes place only in the rhombic modification which is usually called Rochelle salt. The ferroelectric transition is accompanied by a lowering of symmetry from rhombic to monoclinic and by a small monoclinic deformation (at 11° C the monoclinic shear is 3′ and the monoclinic angle $\beta = 90°\ 3'$). The axes of the rhombic (o) and the monoclinic (m) cells are connected by the relations

$$a_o = b_m, \quad b_o = a_m, \quad c_o = c_m.$$

The ferroelectric axis is the monoclinic axis $b_m = a_o$. The space group of the crystal changes in the transition

$$P2_1\,2_1\,2_1 \rightleftarrows P2_1.$$

The determination of the structure of Rochelle salt particularly in the ferroelectric region is a very difficult problem, for the construction of only one two-dimensional projection requires a knowledge of 112 atomic coordination parameters. The addition of anisotropy of thermal vibrations still further complicates the problem. The difficulties are aggravated by the smallness of the structural and thermal effects in Rochelle salt as compared with the analogous effects in barium titanate or ammonium phosphate.

Only neutron spectroscopy using the method of deuteration has enabled the true symmetry of the ferroelectric modification to be obtained. The neutron scattering amplitudes of the hydrogen isotopes differ in sign ($H = -0{\cdot}38 \times 10^{-12}$ cm; $D = 0{\cdot}65 \times 10^{-12}$ cm), so crystallising from a solution which has a suitable concentration of heavy and light water, we can obtain partially deuterated crystals of Rochelle salt in which the neutron scattering of the H and D atoms cancel each other out. The absence of the odd orders of reflection $0k0$ in the ferroelectric phase indicate a $P2_1$ group. These results show for certain that the spontaneous polarisation is connected with displacements of the hydrogen atoms. Fig. 155 shows the projections of the crystal structure of Rochelle salt as obtained by X-ray and neutron diffraction; the lower projection gives the positions of all the atoms including those of hydrogen.

The Domain Structure of Ferroelectric Crystals

Below the Curie point, ferroelectrics even in single crystal form are similar to ferromagnetics in that, in the absence of an external field, there is no apparent macroscopic spontaneous polarisation. This is explained by the existence of regions in the crystal—the domains—with different directions of spontaneous polarisation, as a result of which the polarisation vectors of the individual domains in the crystal cancel each other out and the resultant polarisation is zero or very small, although the individual domains are strongly polarised. The reason for the appearance of domains, as was first shown by Landau and Livshits for ferromagnetics, is that the polarisation energy of the crystal is thereby lowered. The energy of polarisation is equal to the

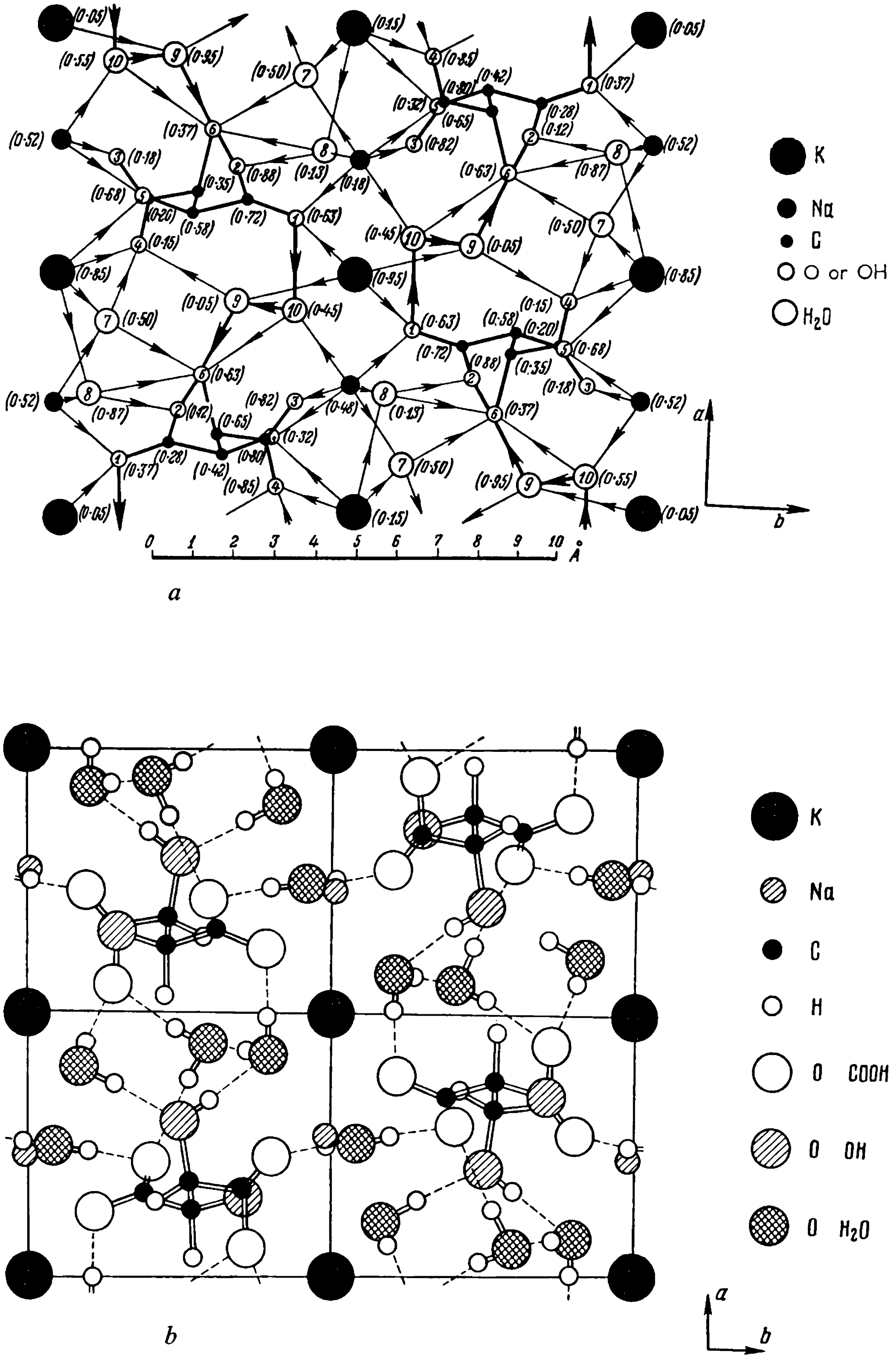


FIG. 155. Projections of the crystal structure of Rochelle salt: (*a*) as obtained from the X-ray results of Beevers and Hughes, (*b*) as obtained from the neutron diffraction results of Frazer.

product of the depolarising field and the polarisation. When domain structure is formed, the crystal surface becomes a mosaic of positively-charged, negatively-charged and neutral regions, and the mean surface charge and the depolarising field become zero.

5. Electrical Conductivity of Crystals

Electrical Conductivity of Solid Electrolytes

The mobility of ions in a crystal under the influence of an electric field is closely connected with their diffusional mobility. This is shown by the fact that, at sufficiently high temperatures, the temperature dependence of the electrical conductivity of solid

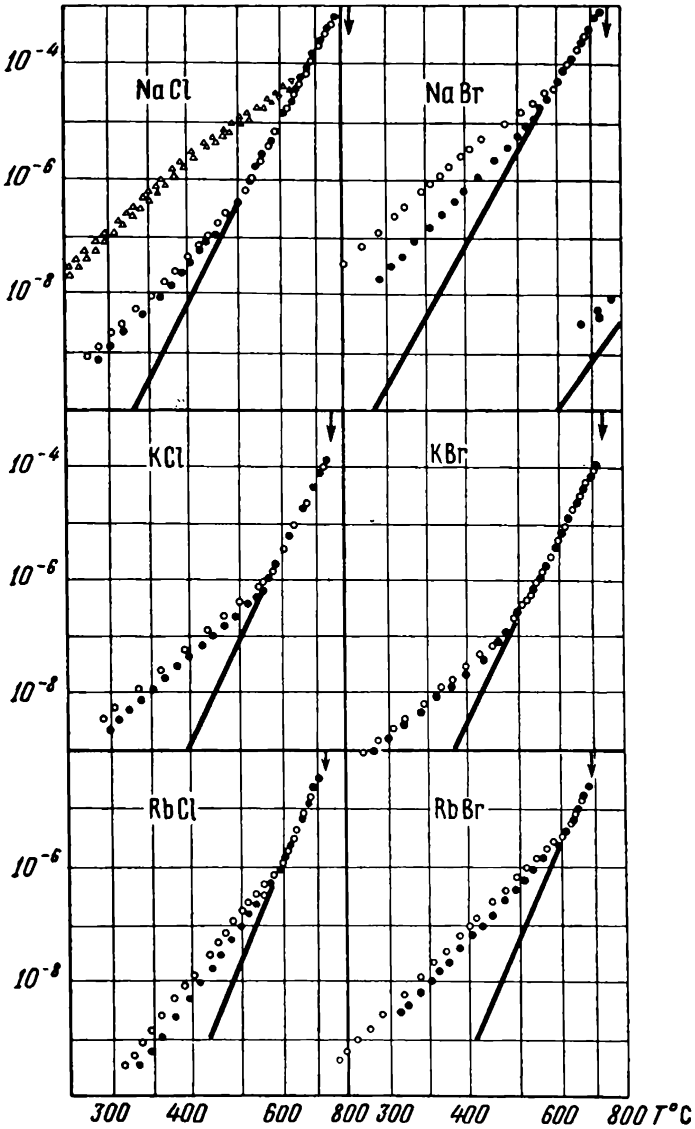


FIG. 156. Temperature dependence of the electrical conductivity of alkali halide compounds.

electrolytes as well as of the coefficient of diffusion are expressed by the exponential formula

$$\sigma = A \exp(-Q/kT). \qquad \ldots [9.112]$$

Fig. 156 shows the dependence of the logarithm of the electrical conductivity on $1/T$ for a variety of alkali halide compounds. At high temperatures the experimental points for two artificially grown pure single crystals of each substance are in satisfactory coincidence with the straight lines which correspond to formula [*112*]. At low temperatures the electrical conductivity depends mainly on the history of the specimen and its impurity content. The marked divergence in the results for NaCl is explained by the fact that the measurements were carried out, not with artificial, but with natural crystals of rock salt.

The mobility of the ions depends on the ionic radius, charge, and so on, and can be very different even for the ions in a single crystal. Table 59 gives the ionic transport numbers, that is the fraction of ions of each kind in the current. As the Table shows, in the halides of silver and the alkali metals at not too high temperatures, mobility is a property of the cations because they have smaller ionic radii than the anions.

Table 59

Transport Numbers for Solid Electrolytes at Various Temperatures

Compound	Temperature in °C	n_+	n_-	Compound	Temperature in °C	n_+	n_-
NaF	500	1·000	0·000	AgCl	20–350	1·00	0·00
	550	0·996	0·004	AgBr	20–300	1·00	0·00
	600	0·916	0·084	BaF_2	500	0·00	1·00
	625	0·861	0·139	$BaCl_2$	400–700	0·00	1·00
NaCl	400	1·000	0·000	$BaBr_2$	350–450	0·00	1·00
	510	0·981	0·019	PbF_2	200	0·00	1·00
	600	0·946	0·054	$PbCl_2$	200–450	0·00	1·00
	625	0·929	0·071	$PbBr_2$	250–365	0·00	1·00
KCl	435	0·956	0·044	PbI_2	255	0·39	0·61
	500	0·941	0·059		290	0·67	0·33
	550	0·917	0·083				
	600	0·884	0·166				

In the halides of barium and lead on the other hand the anions possess the mobility. As the temperature is raised there is a rise in mobility and different ions may take part in the current at the same time.

In the electrolytes, oppositely charged ions move in opposite directions along the lines of force of the field. As a result of this, ionic conductivity is accompanied by electrolysis, that is, by chemical decomposition of the substance under the influence of the electric field. Electrolysis obeys the well-known Faraday's Law, which states that the amount of a substance deposited at an electrode is exactly determined by the total electric charge flowing through the electrolyte and by its electro-chemical equivalent. If a part of the electric current is carried by electrons we can determine the fractions of ionic and electron conductivity by comparing the electrical and material currents.

Electrical Conductivity of Metals

It is well known that metals are good electrical conductors and that the current carriers are electrons. The latter can be demonstrated by a variety of experiments: for example Riecke, aiming to demonstrate electrolysis in metals, passed a constant electric current for ten years through a cylinder composed successively of copper, aluminium and copper. The result of this experiment was that there were no measurable traces of electrolysis, and shows the electronic nature of metallic conductivity. The nature of the current carriers in conductors can also be established by determining the ratio e/m for the current carriers by means of electron inertia experiments. The idea behind such an experiment is that, if the motion of a piece of metal is accelerated (either by an impulse or by sudden stopping) the free electrons will lag behind or get ahead of the atomic lattice. This produces a rapid redistribution of the electric charges which leads to the appearance of some potential difference which may show up as a current pulse. Such an experiment was carried out by Mandelshtam in 1913-14.

The idea of the converse electro-inertial experiment was suggested by Maxwell and was as follows: if an electric current is flowing through a freely suspended metallic coil screened from magnetic fields, then the coil must rotate in one direction when the current is switched on and in the opposite direction when the current is switched off. This experiment was first attempted by Malinowski and repeated with positive results by Barnett (1931). The ratio e/m found by it was close to the corresponding value for cathode rays.

The strength of the current in a metal is linearly related to the applied potential difference (Ohm's Law), and this is applicable within a wide range of current density. Departures from the linear dependence were not observed at current densities up to 10^7 amp/cm^2, which corresponds to the passage of 6×10^{25} electrons per second through each sq. cm of cross-section of the conductor. Experiments with still higher current densities are made very difficult by intense Joule heating of the conductor.

The ability of various conductors to pass an electric current is measured by the coefficient of proportionality in Ohm's Law. In ordinary calculations one frequently makes use of the specific conductivity σ or its reciprocal, the specific resistivity ρ. Both these quantities refer to a volume of metal having the shape of a cube with edge 1 cm, in which volume there are different numbers of atoms for different metals. Theoretically it is more convenient to use the atomic electrical conductivity which is referred to a gm-atom of the substance, also in the form of a perfect cube. If A is the atomic or molecular weight and δ is the density then the volume of the gm-atom is A/δ. Consequently the atomic conductivity is

$$\sigma_A = \sigma\,(A/\delta)^{\frac{1}{3}}. \qquad \ldots [9.113]$$

Table VI of Appendix I gives the specific electrical resistivity of metals at 0° C. As these data show there is no simple relation between ρ and the position of the element in the periodic table (i.e. between ρ and atomic number or basic valency). Nonetheless, there must be a relation between the electrical conductivity and the electron structure of the metal. Thus the highest electrical conductivity is found in monovalent metals, that is metals of the copper group and the alkali metals. As the valency is increased there is a significant reduction in electrical conductivity, as can be seen by comparing the electrical conductivities of the alkali and the alkaline earth metals and also by comparing the metals of the copper and zinc groups. The electrical conductivity of the transition metals is also much less than that of the monovalent metals.

In cubic crystals the electrical conductivity does not depend on the crystal direction. In crystals with other types of symmetry there is anisotropy in the electrical conductivity. In uniaxial crystals the magnitude of the specific electrical resistivity in a direction at an angle ϕ to the principal axis of the crystal is

$$\rho_\phi = \rho_\parallel \cos^2 \phi + \rho_\perp \sin^2 \phi,$$

where $\rho_\parallel$ and $\rho_\perp$ are the magnitudes of the specific resistivity corresponding to electric currents parallel and perpendicular to the principal axis. Table 60 gives data for the anisotropy of the electrical resistivity of several uniaxial crystals.

Table 60

Specific Resistivity (ohm.cm) for Several Uniaxial Crystals

Crystal	System	$t°$ C	$\rho_\parallel \times 10^6$	$\rho_\perp \times 10^6$
Mg	Hexagonal	18	3·37	4·54
Zn	Hexagonal	0	5·83	5·39
Cd	Hexagonal	0	7·65	6·26
β-Sn	Tetragonal	0	13·13	9·05
Sb	Rhombohedral	0	31·8	38·6
Bi	Rhombohedral	0	127	100

In zinc and cadmium crystals, which have structures of the h.c.p. type but somewhat extended along the principal axis, the specific resistivity along the principal axis is greater than in the perpendicular directions. In a magnesium crystal which has a structure somewhat compressed along the principal axis the resistivity along this axis is less than in the other directions. However such a simple relation between the resistivity and the density of packing of atoms is not found in the isomorphic crystals, antimony and bismuth. In bismuth the electrical resistivity in a direction perpendicular to the atomic layers is greater than in directions parallel to the atomic layers. In antimony the relationship between the electrical resistivity in these directions is the opposite of that for bismuth.

The electrical resistivity of disordered solid solutions increases with increasing concentration and does not depend on whether the solute has greater or less electrical resistivity than the solvent (Fig. 157). When order is produced in a solid solution the electrical resistivity falls. Thus minima occur in the curve of electrical resistivity of annealed alloys, these minima corresponding to the formation of ordered structures such as Cu_3Au and CuAu.

Plastic deformation of crystals produces some distortion of the crystal lattice and raises the specific resistivity by amounts which may be as much as several tenths of the resistivity at normal temperatures. In tungsten for example, the increase reaches 50%.

Raising the temperature of a metal crystal increases its electrical resistivity. At not too low temperatures the electrical resistivity increases proportional to temperature. When metals melt there is an increase in electrical resistivity, usually by about a factor of two, but in mercury and in other cases where melting is accompanied by positive changes in volume the increase is about four times. In metals, where the volume is reduced on melting there is quite a large reduction in electrical resistivity, for example, gallium by 53%, antimony by 29%, and bismuth by 54%. At low temperatures there is a departure from the linear dependence and the electrical resis-

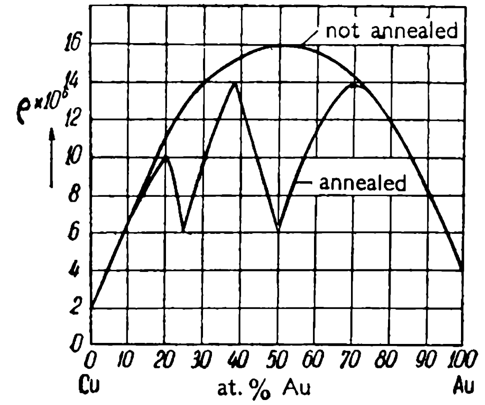


FIG. 157. Electrical resistivity of gold-copper solid solutions in the annealed and non-annealed conditions.

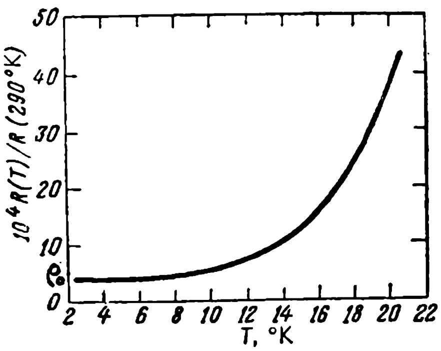


FIG. 158
Electrical resistivity of sodium at low temperatures.

tivity tends to some constant value ρ_0 for a given specimen (Fig. 158). ρ_0 is called the residual resistivity and for a given metal is different for different specimens since it depends on the impurity content and the damage to the structure which has been produced by plastic deformation. The purer the metal and the less it has been deformed, the less is the residual resistivity. The increase in electrical resistivity of metals resulting from damage to the regular crystalline lattice, whether by formation of solid solutions or by plastic deformation or by rise in temperature, shows that a possible reason for the appearance of electrical resistivity in metals may be distortions of the regular periodic structure. If this is so, then we may suppose that the resistivity of pure, undeformed metal crystals will tend to zero when the temperature tends towards 0° K.

The emission of Joule heat when an electric current passes through a metal is evidence for an interaction of the conduction electrons with the crystal lattice in the form of inelastic collisions of electrons with the atomic framework. If a metal crystal is heated unequally, the free electrons take part in the transport of thermal energy from the hotter parts to the cooler parts. If the heat transfer by the free electrons is much greater than the heat transfer by the atoms, then we may expect a relation between the electrical and thermal conductivities of metals. This relation in fact exists and is expressed by the Wiedemann-Franz Law which was originally discovered experimentally

$$\kappa/\sigma = aT. \qquad \ldots [9.114]$$

This shows that the ratio of the conductivities is proportional to the absolute temperature.

6. Superconductivity

Zero Resistivity and the Temperature of Transition to the Superconducting State

If we extrapolate to $T = 0$ we might expect that the dependence of the resistivity of a metal $\rho(T)$ on temperature would be of one of the forms shown in Fig. 159: (Curve

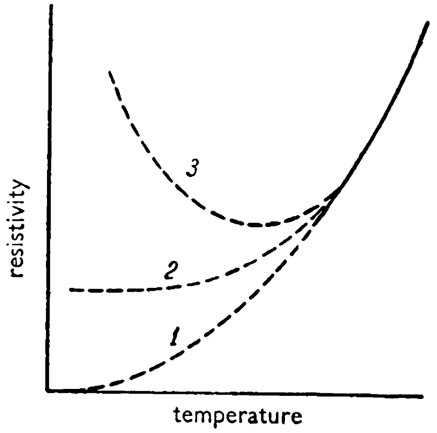


FIG. 159. Possible forms of the dependence of resistivity of a metal on temperature.

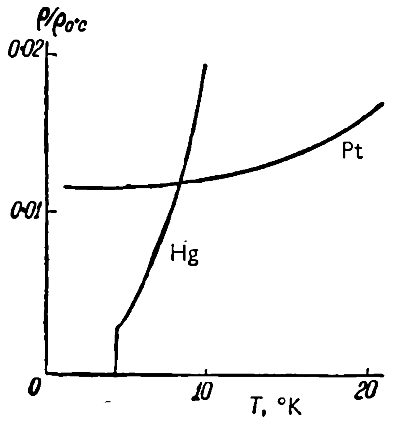


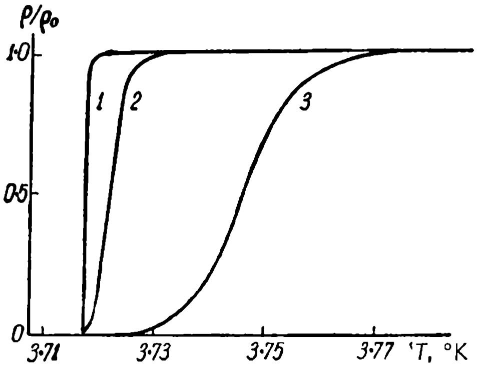
FIG. 160. Dependence of the resistivity of platinum and mercury on temperature.

1), the resistivity is produced by scattering of electrons by thermal atomic vibrations of the crystal and falls to zero; (Curve 2), in the presence of impurities and of lattice deformation the resistivity tends to a constant value; (Curve 3), as a result of reduction in the number of current carriers or of their mobility, the resistivity passes through a minimum.

Experiments to decide this problem were carried out in 1908 using low temperature liquid helium techniques and the results were unexpected: Fig. 160 shows the measured dependence of the resistivity of pure platinum where the curve of $\rho(T)$ tends to some finite value ρ_0. In spite of all attempts to reduce the impurity content, the residual resistivity of platinum remained at this value ρ_0. Kammerlingh Onnes (1911) carried out experiments with mercury which can be obtained in a very pure form and he discovered that at temperature $T_c \approx 4{\cdot}2°$ K the electrical resistivity falls discontinuously and becomes zero. The temperature T_c thus divides the regions of the normal and superconducting states of a metal. The temperature dependence of ρ for mercury thus does not correspond to one of the three cases considered above for the extrapolation of the high temperature part of the curve $\rho(T)$. Another unexpected result was that addition of impurities to the mercury did not suppress appearance of superconductivity and the complete disappearance of resistivity. In a ring made of a superconductor, a constant electric current did not die out in prolonged experiments. Attempts to estimate the decay of the current from these experimental data show that the residual resistivity of a superconductor ρ_s, if it exists at all, must be less than $10^{-17}\,\rho$, where ρ is the resistivity at room temperature. Theory and experiment lead

FIG. 161

Influence of the quality of the specimen on the character of the transition to superconducting state in tin; 1, single crystal; 2, polycrystal; 3, polycrystal with impurities.



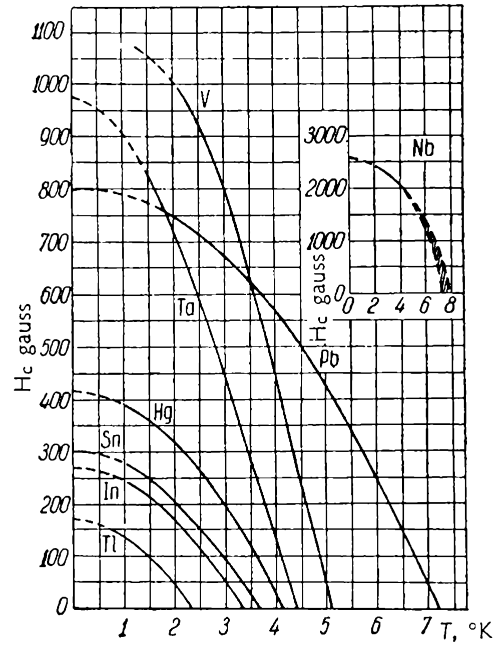


Fig. 162
The dependence of the critical magnetic field strength on temperature.

to the conclusion that $\rho_s = 0$. Fig. 161 shows curves of the transition to the superconducting state for single and polycrystals of pure tin and also for polycrystalline tin with impurities. These show that the homogeneity and purity of the material have a marked influence on the sharpness of the transition.

The disappearance of resistivity of a metal at a finite temperature when it still contains dynamic lattice defects, and the existence of superconductivity in an impure metal with static stresses are similar phenomena as regards scattering of electron waves. That is, lattice distortions of one kind or another do not prevent the appearance of superconductivity. This shows that the conduction electrons in the superconducting state are transferred to a special energy state and do not interact with the crystal lattice.

The Influence of a Magnetic Field

A sufficiently strong magnetic field destroys the superconducting state. The magnitude of the field H_c required for the transition from the superconducting to the normal state, depends on temperature and increases as T falls (Fig. 162). For many semiconductors this dependence can with sufficient accuracy be described by a parabola

$$H_c = H_0\,[1-T/T_c)^2], \qquad \ldots [9.115]$$

where H_0 is the critical field for complete destruction of superconductivity at temperatures such that $T \to 0$. The superconductivity can also be destroyed by the influence of the magnetic field produced by the current itself. Thus there is a limiting current strength in a superconductor. For a long straight wire of radius r the limiting current is determined by the relation

$$H_c = 2I/r. \qquad \ldots [9.116]$$

that is, by the condition that a critical field strength is reached on the surface of the wire. If the current exceeds the value given by formula [*116*] then ohmic resistance is produced in the metal and Joule heat begins to be evolved. As experiments with lead wires show, this leads to melting of the wire even if it is immersed in liquid helium. Such melting is explained by bad heat transfer resulting from evaporation of helium and the development of gaseous bubbles on the surface of the wire.

The Influence of Pressure

An external pressure p produces changes in T_c and changes in the strength of the magnetic field required to destroy superconductivity. Thus in tin and indium there is an increase in H_c and $\delta H_c/\delta p \sim 10^{-8}$. The magnitude of the change depends on temperature. For tantalum we have

T	3·90	4·00	4·11	4·22
$\delta H_c/\delta p$	1·9	2·6	3·8	$4{\cdot}7 \times 10^{-9}$

In the case of thallium one finds a non-linear dependence (Fig. 163): at low pressures T_c increases, then passes through a maximum and at high pressures falls again.

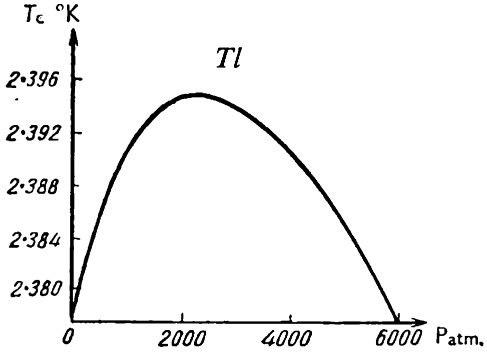


FIG. 163. The influence of pressure on T_c in thallium.

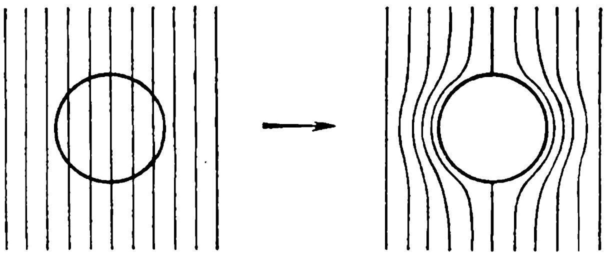


FIG. 164. Ejection of the flux of magnetic induction from a superconductor.

Ejection of the Magnetic Flux (Meissner, 1933)

When a superconductor is cooled in a magnetic field then at the transition to the superconducting state the lines of magnetic induction are pushed out of it (Fig. 164). The superconductor behaves as an ideal diamagnetic, inside which B = 0. Thus for the diamagnetic susceptibility of a semi-conductor we have

$$\chi_{\text{superconducting}} = -\tfrac{1}{4}\pi.$$

This property of a superconductor does not arise as a consequence of the destruction of resistivity, for we can show that an ideal conductor is not an ideal diamagnetic. From Ohm's Law $E = \rho j$ it follows that at a finite current density j, if $\rho = 0$ then $E = 0$ and thence also curl $E = 0$.

Maxwell's equation,

$$\frac{dB}{dt} = -c \operatorname{curl} E, \qquad \ldots [9.117]$$

then shows that $dB/dt = 0$ and that consequently the flux of magnetic induction does not change in an ideal conductor.

Influence of the Shape of the Specimen. The Demagnetising Factor

The shape of the specimen has an important effect on magnetisation: only for an infinitely long cylinder magnetised parallel to its axis are all points on the surface subject to a field of uniform intensity, and in this case the magnetisation curve increases linearly with magnetic field strength (Fig. 165*a*). When the value of the critical field H_c is reached the magnetisation falls sharply; the metal changes into the normal state and loses its property of being an ideal diamagnetic.

Fig. 165*b* shows the magnetisation curve of a superconducting sphere. Here a linear increase of magnetisation takes place until the external field is $H = \frac{2}{3} H_c$ when the resultant magnetic field on the equator of the sphere reaches the critical value and

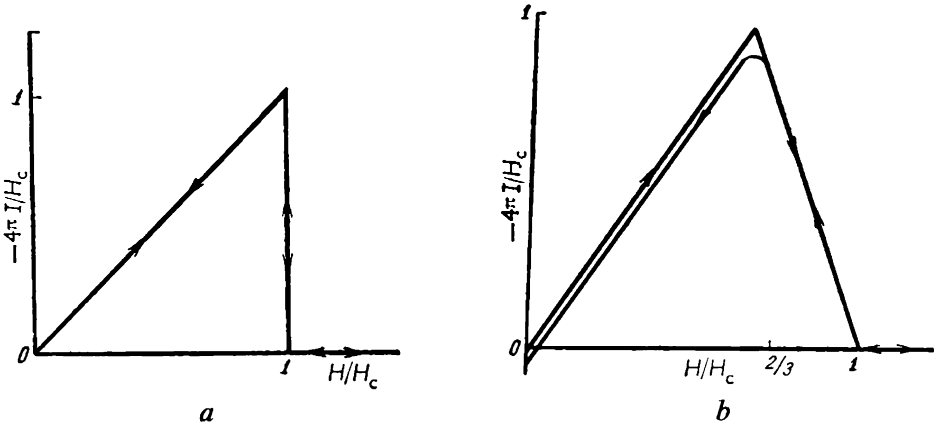


FIG. 165. Magnetisation curves: for (*a*) a superconducting long cylinder, (*b*) a superconducting sphere.

the equatorial region begins to transform to the normal state. As the external field is increased the volume of the superconducting part decreases and at $H = H_c$ the transition to the normal state is complete. At field strengths which lie between $\frac{2}{3} H_c$ and H_c the sphere is in an intermediate state.

The Intermediate State

The transition from the normal to the superconducting state is a phase transition. The intermediate state is a heterogeneous mixture of the superconducting and normal phases and the boundary dividing the superconducting and the normal regions has a surface energy. Landau showed theoretically that the intermediate state of a superconductor must be a complex branched system of layers of both phases. Its topography was investigated experimentally by Shal'nikov and Meshkovsky using a tin sphere with diameter 40 mm consisting of two hemispheres between which there was a plane gap of thickness 0·2 mm. A very small bismuth spiral was introduced into this gap, and the magnetic field in different parts of the gap determined by measuring the resistance of the spiral. Fig. 166 shows the map of the superconducting and normal regions thus obtained.

The Specific Heat Discontinuity

As the solid line in Fig. 167 (Keesom and van Laer) shows, there is a specific heat discontinuity at the transition to the superconducting state in tin. This discontinuity is analogous to that observed when a system passes from the disordered to the ordered state and shows that the superconducting transition is a second order phase transition. Near the transition point the entropy of a superconductor increases relative to that

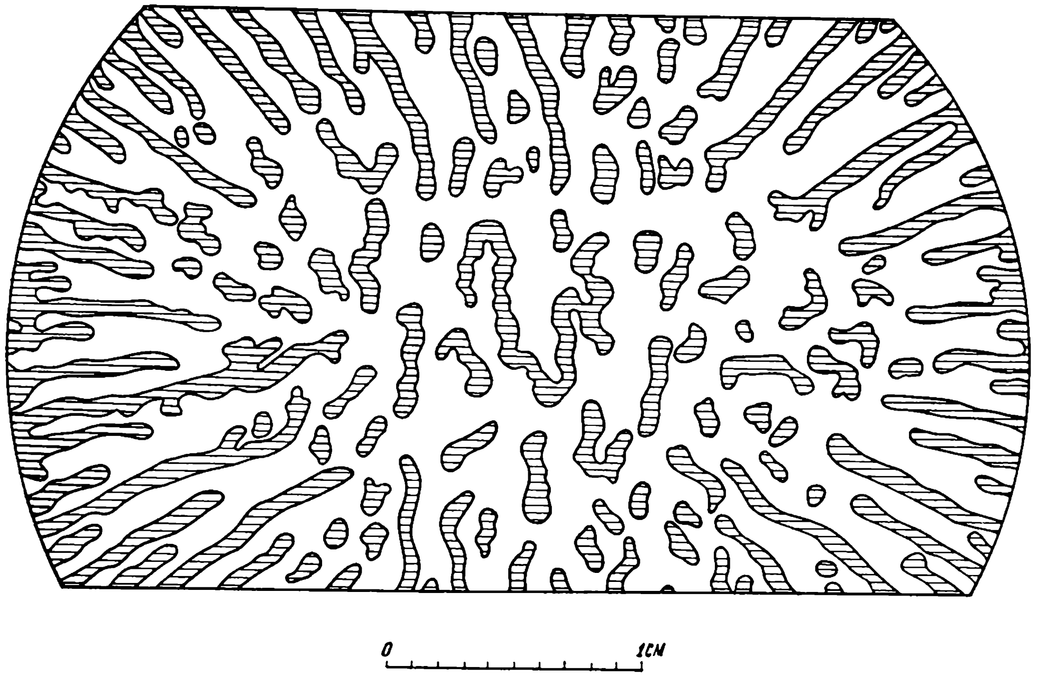


FIG. 166. Topography of the intermediate state formed when an external field is applied in an equatorial gap between two tin hemispheres.

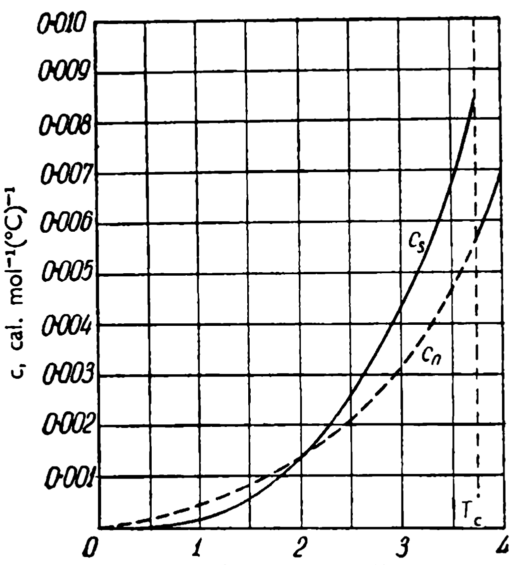


FIG. 167. The discontinuity in thermal conductivity in tin at the transition from the normal (C_n) to the superconducting state (C_s).

of the normal state. In absolute magnitude the specific heat discontinuity is small and for tin is

$$\Delta c = 0{\cdot}003 \text{ cal/mole . °C},$$

which is an order of magnitude less than the discontinuity at the magnetic transition

$$(\Delta c = 0{\cdot}03 \text{ cal/mole . °C in nickel})$$

and almost two orders of magnitude less than the discontinuity arising from atomic ordering (Fig. 209)

$$(\Delta c = 0{\cdot}15 \text{ cal/mole . °C in } \beta\text{-CuZn}).$$

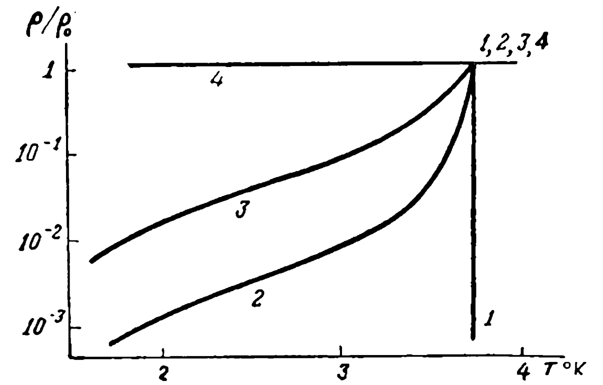


FIG. 168. The temperature dependence of the resistivity of tin at various frequencies.

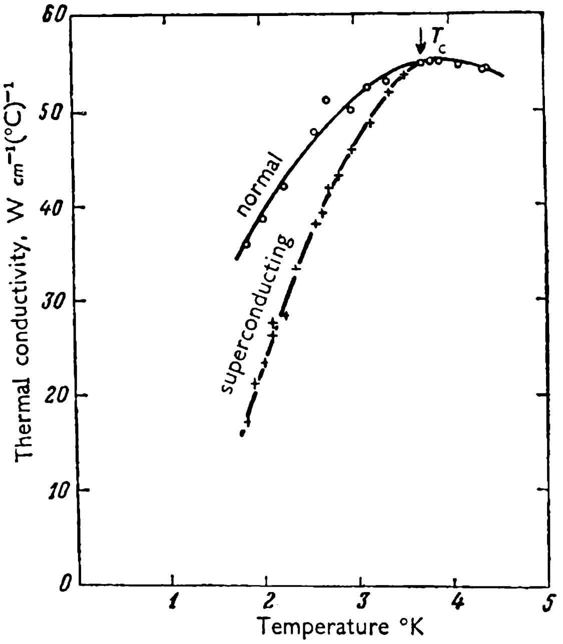


FIG. 169. The thermal conductivity of tin in the superconducting and normal states.

This indicates that the interaction which produces the transition to the superconducting state is small. The dotted curve in Fig. 167 gives the specific heat for the normal phase measured in a magnetic field strong enough to destroy superconductivity, and this curve is the extrapolation of the curve $c(T)$ above T_c. Destruction of superconductivity in a magnetic field is accompanied by the emission of the latent heat of the transition and is a first order phase transition.

The Gyromagnetic Ratio

Kikoin and Gubar measured the gyromagnetic ratio for a superconductor and found it to be $-e/2mc$, which corresponds to the ratio for orbital moments. Consequently, diamagnetism of a superconductor is produced by electric currents and is not connected with the electron spins.

The Frequency Dependence

Fig. 168 shows the temperature dependence of the resistivity of tin at various alternating current frequencies. For a constant current and for frequencies less than 10^7 c/s there is an abrupt fall of the resistivity to zero (curve 1). However at higher frequencies the resistivity of a superconductor below T_c remains at a finite value which is larger the higher is the frequency (curves 2 and 3). In the infra-red region ($\sim 2 \times 10^{13}$ c/s)

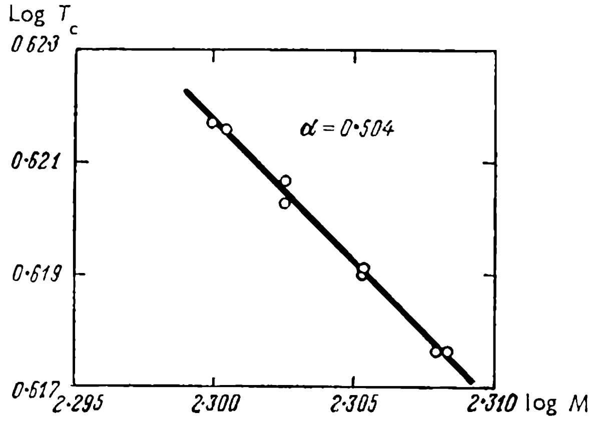


FIG. 170
Dependence of the logarithm of the critical temperature on the logarithm of the mean value of the mass number for various isotopic mixtures of mercury.

the absorption in the normal and in the superconducting states is the same and does not depend on temperature (curve 4).

Thermal Conductivity

In pure metals, the transition to the superconducting state is accompanied by a reduction in thermal conductivity. This can be seen for tin in Fig. 169 and for lead in Fig. 247 (page 406), and shows that superconducting electrons cease to interact with the lattice and do not take part in heat transport.

The Isotope Effect

Fig. 170 shows, on a logarithmic scale, the critical temperature for specimens of mercury with different isotopic composition. Increase of the mass number M from 199·5 to 203·4 leads to a reduction of the temperature T_c from 4·185 to 4·146° K. The experimental points on Fig. 170 fall very well on a straight line with slope close to ½. Thus

$$M^{\frac{1}{2}}T_c = \text{const.} \qquad \ldots [9.118]$$

Relation [*118*] shows the connection of superconductivity with the maximum frequency in the thermal vibration spectrum of the crystal ν_m, or with the characteristic temperature θ. As formula [*14.28*] shows $h\nu_m = k\theta$ (we leave out the index m on the θ) and from [*14.50*] it follows that the characteristic temperature is proportional to the velocity of sound c. Now according to [*14.30*] the velocity of sound is reciprocally proportional to the square root of the density, and the density of a body under given conditions is proportional to the mass of the particles composing it, M. Hence it follows that the critical temperature of a superconductor under uniform conditions is proportional to the characteristic temperature or to the maximum energy of phonons in the spectrum of free oscillations of the crystal.

The Energy Gap. Investigations with Photons and Phonons

The transition from the normal to the superconducting state is a sharp second order phase transition and is accompanied by small changes in other properties. Thus, in a transition through T_c, the velocity of sound (at $\omega = 5 \times 10^4$ c/s) in tin, changes by a few millionths. There are also changes of state of the valence electrons, which below T_c do not interact with the lattice and this leads to a change of specific heat, thermal conductivity, thermoelectric properties and so on. The superconducting (ground) state is separated from the normal (excited) state below T_c by an energy gap $\Delta E_s \sim 10^{-4}$ eV.

The first indications of the existence of the energy gap were obtained from the exponential law for the reduction in the electron specific heat in a superconductor

$$c_{es} \sim \gamma T_c \exp(-bT_c/T) \sim c_{ns} \exp(-bT_c/T).$$

The existence of such a gap can be demonstrated by various delicate methods in which the means of investigation are photons or phonons. Thus, the transition of electrons from the superconducting level to the normal level must be accompanied by absorption of electromagnetic waves. The absorption band lies in the region $h\nu \sim kT_c$, which corresponds to a frequency of 30-150 kilomegacycles per second, or a wavelength of 1 cm or less. The effect of this absorption at a temperature below

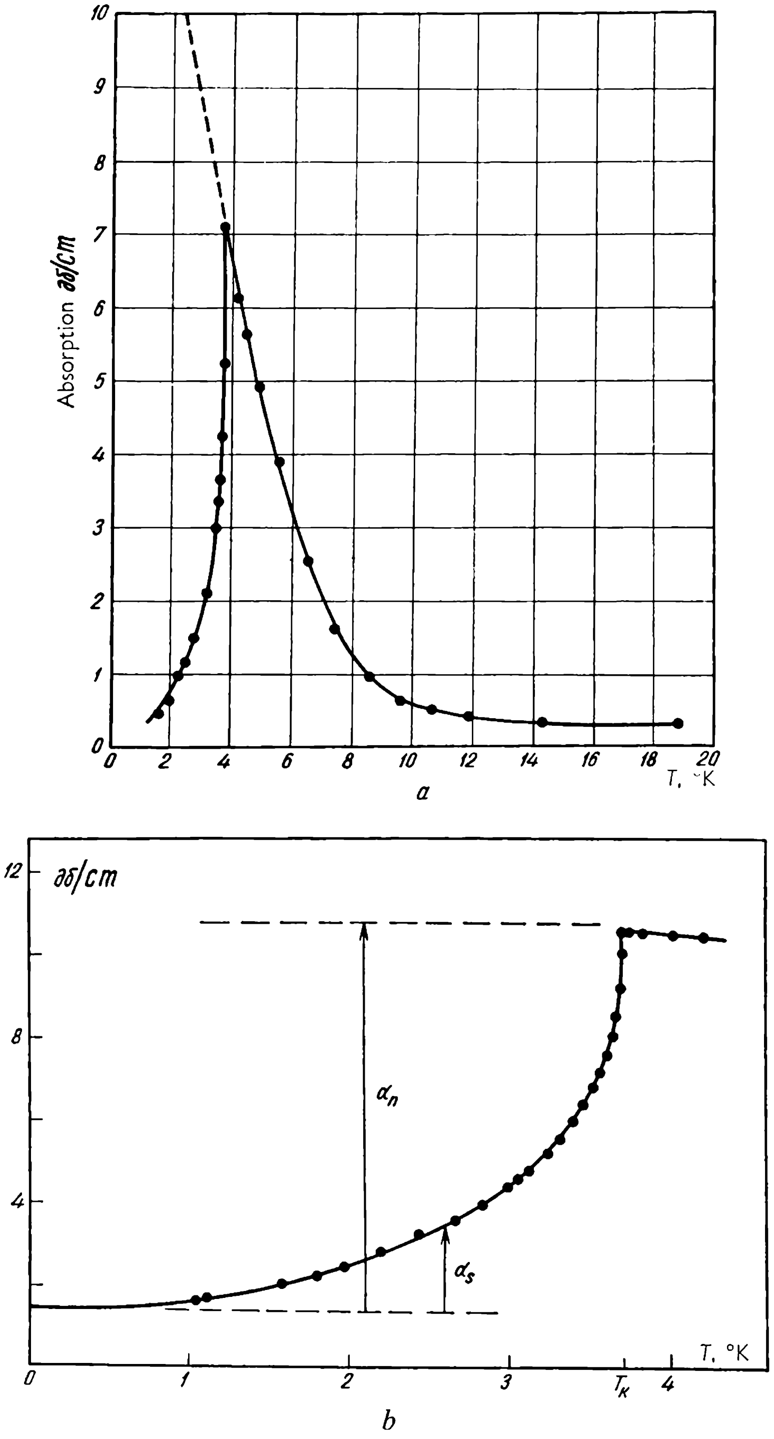


FIG. 171. The temperature dependence of the absorption of supersonic waves in a single crystal of tin: (*a*) for ω = 10·3 Mc/s; the dotted line shows absorption for the normal state; (*b*) for ω = 33·5 Mc/s.

the critical can actually be observed in tin and aluminium when the width of the gap $\sim 2\ kT_c$.

In pure metals at temperatures below 10° K there is a sharp increase of ultrasonic absorption in the range 10-40 Mc/s (Fig. 171) which is explained by the absence of interaction between electrons and lattice in the superconducting region leading to a greatly reduced damping. This effect is observed in In, Sn and Pb, particularly interesting results having been obtained recently by Morse and Bohm who were carrying out detailed investigations of damping in the superconducting region (Fig. 171*b*). According to the microscopic theory, the ratio of the damping coefficients in the superconducting and normal phases is given by the formula

$$\alpha_s/\alpha_n = 2/(\exp\{\Delta E_s/kT\}+1), \qquad \ldots [9.119]$$

where ΔE_s is the width of the energy gap. Fig. 172 shows the change of width of the gap $\Delta E_s(T)$ as a function of temperature; from the experimental data $\Delta E_s(0) = 1{\cdot}77\, kT_c$ while theory gives 1·75 for the numerical coefficient.

Thermodynamics of Superconductivity

For the sake of simplicity we shall leave out those complications which are introduced by the shape of the specimen and are connected with the intermediate state. Suppose that the specimen has the shape of a long cylinder and is situated in a magnetic field H parallel to the cylinder axis. Take the temperature T as constant and less than T_c. A sufficiently strong field destroys superconductivity and transfers the specimen to the normal state. Call the free energy of the normal state F_n. When the strength of the magnetic field is reduced to the critical value H_c at a given temperature T, the specimen goes over to the superconducting state. In this state, call the free energy F_s. Now let us find the difference of the free energies of these states. Neglect volume changes in the metal in the superconducting transition and also

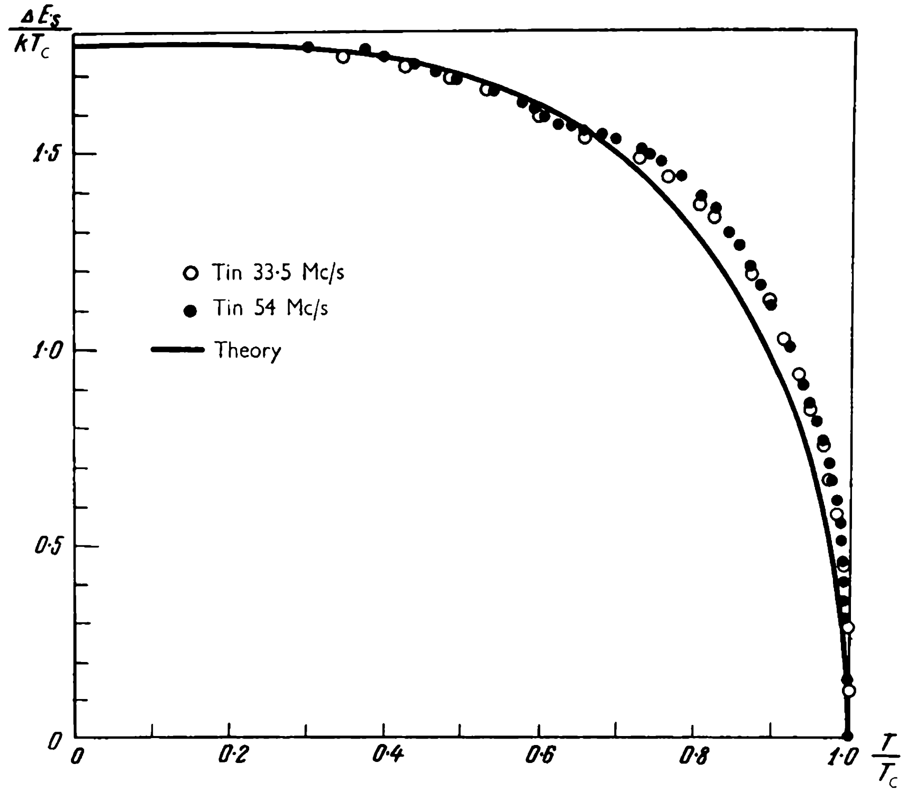


FIG. 172. The temperature dependence of the width of the gap ΔE_s between the normal and the superconducting states of tin; the unbroken curve is drawn from equation *9.119* with $\Delta E_s(0) = 1{\cdot}77\, kT_c$.

neglect the surface energy on the boundary of a superconducting phase. In this approximation, the change of energy in the transition to the superconducting state is determined by the Meissner effect and is equal to the energy of the magnetic field which is driven out from the superconductor. Consequently

$$F_n - F_s = H_c^2/8\pi. \qquad \ldots [9.120]$$

Formula [*120*] is basic to the thermodynamics of phase transitions from the normal to the superconducting state.

The entropy $S = -\frac{dF}{dT}$, so difference in entropy between the normal and superconducting phases is

$$S_n - S_s = -(H_c/4\pi)\frac{dH_c}{dT} \qquad \ldots [9.121]$$

The differential coefficient $\frac{dH_c}{dT}$ is always negative so the entropy of the superconducting phase is less than or equal to the entropy of the normal phase. At the transition point, $H_c = 0$ so the entropies are equal at this point. Isothermal destruction of superconductivity by a magnetic field is accompanied by the absorption of heat $q = T(S_n - S_s)$. Substituting from [*121*] we obtain

$$q = -(TH_c/4\pi)\frac{dH_c}{dT}. \qquad \ldots [9.122]$$

In the absence of a magnetic field ($H_c = 0$) the heat of transition $q = 0$ since at T_c the differential coefficient $\frac{dH_c}{dT}$ remains finite (cf. [*115*]), and so the transition to the superconducting state in this case is of second order. In a magnetic field this transition is accompanied by the absorption of heat. In the reverse direction there is emission of heat and it is thus a first order phase transition.

The specific heat is $C = T\left(\frac{dS}{dT}\right)$ and the difference of specific heat between the superconducting and the normal phase is thus:

$$\Delta c = \frac{T_c}{4\pi}\left(\frac{dH_c}{dT}\right)^2 + \frac{TH_c}{4\pi}\frac{d^2H_c}{dT^2}. \qquad \ldots [9.123]$$

In the absence of a magnetic field, that is when $T = T_c$ we obtain Rutger's formula

$$\Delta c = \frac{T_c}{4\pi}\left(\frac{dH_c}{dT}\right)^2_{T=Tc}, \qquad \ldots [9.124]$$

from which it follows that at the transition temperature the specific heat changes discontinuously.

Comparison with Experiment

We have already given experimental data on the specific heat discontinuity in the superconducting transition of tin in the absence of a field and on the existence of a latent heat of transition in a magnetic field. These show experiments are in qualitative agreement with the conclusions of the thermodynamic theory. That there is also good quantitative agreement with the values calculated by the thermodynamic formula [*124*] is shown by the experimental data in Table 61.

Table 61

Comparison of Formula (124) with Atomic Heat Measurements

Metal	T_c °K	Atomic Volume V cm^3/gm. at.	$(dH_c/dT)_{T=T_c}$ gauss/°C	$\Delta c \times 10^3$ cal/°C.gm. at.	
				Calculated	Measured
Pb	7·22	17·8	200	10	12·6
Ta	4·40	10·9	320	9·4	8·2–9
Sn	3·73	16·1	151	2·61	2·4–2·9
In	3·31	15·2	146	2·08	2·3
Tl	2·38	16·8	139	1·47	1·48
Al	1·20	9·9	177	0·71	0·46

New Ideas about the Microscopic Theory

Recently Bardeen, Cooper, Bogolyubov and others (1957) have worked out a microscopic theory of superconductivity which in principle gives hopes of understanding the phenomenon. We shall now briefly consider the physical basis of this theory which applies to an idealised model where, for the present, the structural peculiarities of the metal are completely neglected. Thus the metal is considered as a potential well filled with an electron gas obeying Fermi statistics. Between individual electrons there act forces of Coulomb repulsion weakened in large measure by the field of the atomic lattice. The isotope effect in superconductivity appears through the interaction of the electrons with the thermal vibrations of the lattice, that is, with the phonons.

As Fröhlich pointed out in 1950 forces of attraction may arise between individual electrons as a result of the interaction of the electrons with the lattice. The mechanism of this interaction is analogous to the mechanism of Coulomb interaction between charged particles in a vacuum—that is the particles are supposed to be continuously exchanging virtual phonons. As quantum electrodynamics shows, the result of this exchange is the formation of an interaction between any two particles, the energy of which is proportional to the product of their charges and reciprocally proportional to the distance between them. The sign of the energy is determined by the product of the signs of the charges of the particles. This is Coulomb's Law. Free electrons in a metal lattice lying near the upper limit of the filled energy band, that is in a narrow layer close to the Fermi surface, also exchange photons but this exchange is weakened as was shown above by screening. The important exchange here is of phonons which leads to a weak attraction between the electrons. If this attraction is larger than the Coulomb repulsion then connected electron pairs are formed with lowered energy. The energy of binding of such pairs, Δ, is expressed by a rapidly varying function of the resultant moment k and reaches a maximum when $k = 0$, that is for electrons situated on diametrically opposite ends of the Fermi surface.

Electron pairs in which the spins are compensated have resultant spin zero, and behave like particles obeying Bose statistics. Now a gas of Bose particles has very different properties from a gas of Fermi particles. In particular at sufficiently low temperatures it exhibits superfluid properties, and for a gas of charged particles this state becomes one of superconductivity. The phenomenon of superfluidity in helium was discovered by Kapitza in 1938; of the two helium isotopes, He^4 and He^3, superfluidity is a property of He^4, the atoms of which have even resultant spin, and not of He^3, the atoms of which have uneven spin. The macroscopic theory of superfluidity was developed by Landau (1941) who was working on the hydrodynamics of quantum liquids, while the microscopic theory of superfluidity was put forward by Bogolyubov (1947). There is a great similarity between the phenomena of superfluidity and superconductivity.

At absolute zero the energy level of the electron pairs described above (ground state) lies a distance $\Delta = \omega \exp(-1/\rho)$ below the Fermi level, where ω is the Debye frequency of the crystal spectrum and ρ is a quantity which measures the binding of electrons to the lattice. In the ground state the electron pairs do not interact with the lattice and the metal exhibits superfluidity. Heating the metal is accompanied by breaking up of the electron pairs and transition of electrons from the ground to excited levels which belong to the energy spectrum of a metal obeying the normal Fermi statistics. The maximum temperature of complete transition of the metal to the normal state is the transition temperature at which superconductivity is destroyed $T_c = 1{\cdot}14\theta \exp(-1/\rho)$, where θ is the Debye temperature. Preliminary calculations give $1/\rho > 2$, so the temperature of transition to the superconductivity state is an order of magnitude lower than the Debye temperature.

The microscopic theory of superconductivity has clarified the basic facts about superconductivity. These are the dependence of T_c on the isotopic mass, the infinite conductivity ($E = 0$), the Meissner effect ($B = 0$), the exponential dependence of the electron specific heat near $T = 0$, and the second order phase transition at T_c. Some of the conclusions of the theory show good quantitative agreement with experiment but there are still many unsolved problems, including: the distribution of superconducting metals in the periodic system, the dependence of T_c on the composition and structure of superconducting compounds, the possibility of producing superconductors with the highest possible transition temperature and so on.

Problems

1. Calculate the polarisability of a hydrogen atom in the Bohr approximation.
2. Find the polarisability of a bound electron (frequency of free vibration ω_i) in the field of an electromagnetic wave of frequency ω.
3. Find the depolarising field in the centre of a uniform polarised sphere.
4. Find the value of the internal field for a uniformly polarised system of dipoles.
5. Find $(\cos^2 \phi_i)_{\text{mean}}$ in formula [*1*] of the solution of the previous problem for a cubic system of dipoles.
6. A specimen in the form of a small cube with edges parallel to the axes x_1, x_2, and x_3 is cut out from a crystal of quartz. Find the deformation of the specimen if a difference of potential of a thousand volts is applied to its faces perpendicular to the x-axis. The cube edge is 1 cm.
7. Find the relationship between the components of the piezoelectric tensor of group 2.

Chapter 10

MAGNETIC PROPERTIES OF CRYSTALS

1. Basic Types of Magnet

Magnetization and Magnetic Susceptibility

When a body is brought into a magnetic field it becomes magnetized and a magnetic moment $\boldsymbol{M}$ is developed which is the resultant of the elementary magnetic moments $\boldsymbol{m}_0$ belonging to the individual particles of the body. The magnetization $\boldsymbol{I}$ is defined as the magnetic moment per unit volume

$$\boldsymbol{I} = \boldsymbol{M}/v = \sum \boldsymbol{m}_0 \qquad \ldots [10.1]$$

The summation in formula [*1*] is vectorial. In practice one frequently uses the specific magnetization σ which is related to the unit of mass

$$\sigma = M/m = I/\rho \qquad \ldots [10.2]$$

where m is the mass and ρ the density of the substance. In theoretical work the most useful quantity is the molar magnetization σ_A which relates to the gram atom or to the gram molecule of the substance

$$\sigma_A = (M/m)A, \qquad \ldots [10.3]$$

where A is the atomic or molecular weight. The magnetization increases with increasing intensity of the magnetic field $\boldsymbol{H}$, and corresponding to the three definitions of magnetization [*1-3*] we have

$$\boldsymbol{I} = \kappa \boldsymbol{H}; \quad \sigma = \chi \boldsymbol{H}; \quad \sigma_A = \chi_A \boldsymbol{H}. \qquad \ldots [10.4]$$

The coefficients of proportionality in [*4*] are called the magnetic susceptibilities, κ is the volume susceptibility, χ the mass susceptibility, and χ_A the molar susceptibility. They are related to each other by the equations

$$\chi = \kappa/\rho; \quad \chi_A = (\kappa/\rho)\, A. \qquad \ldots [10.5]$$

Diamagnetism and Paramagnetism

If the magnetic susceptibilities in [*4*] are scalar quantities then the magnetization vector $\boldsymbol{I}$ is collinear with the vector $\boldsymbol{H}$. There are two possible cases: $\boldsymbol{I}$ is (1) parallel, (2) antiparallel to the vector $\boldsymbol{H}$ and these correspond to two classes of substance: paramagnetics and diamagnetics. In a non-uniform magnetic field, paramagnetics set themselves in the region of greatest intensity while diamagnetics move out of the field. For para- and diamagnetic bodies the coefficients κ, χ and χ_A are positive or negative constants, that is the dependence of magnetization on field strength is linear (Fig. 173). In the majority of cases the numerical value of the magnetic susceptibility of diamagnetics is small compared with that of paramagnetics.

Ferromagnetics

When the temperature of some paramagnetics is lowered a magnetic transition takes place, characteristic of a second order phase transition (discontinuity of thermal capacity etc.). The relation between the vectors $\boldsymbol{I}$ and $\boldsymbol{H}$ ceases to be linear and we find a more complex curve (Fig. 174), which becomes a hysteresis loop when the magnetization varies cyclically. Even in weak fields the magnetization is of the order of several gauss and tends to a saturation value, I_s, which is many times greater than the magnetization of common paramagnetics. The state of a body thus described is called ferromagnetism and the temperature of the ferromagnetic transition is called the Curie temperature. A typical example of a ferromagnetic is α-iron, the Curie temperature for which is 780° C. Above the Curie temperature, iron behaves as an ordinary paramagnetic. It should be noted that the temperature of the magnetic transition does not correspond to the temperature of the polymorphic $\alpha \rightleftarrows \gamma$ transition of iron, which is 906° C.

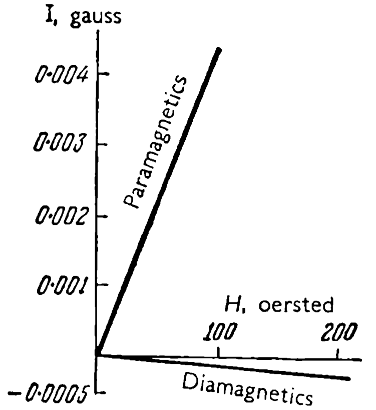


FIG. 173. General form of the dependence of I on H for dia- and para-magnetics.

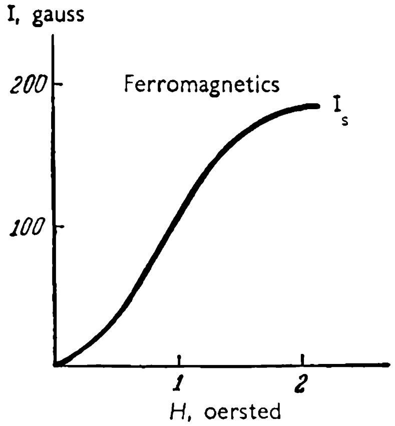


FIG. 174. Dependence of I on H for a ferromagnetic.

Magnetic Permeability

The normal magnetic state of a ferromagnetic is one of spontaneous magnetization. The field inside a ferromagnetic is measured by the magnetic induction $\boldsymbol{B}$ and is made up of the applied field $\boldsymbol{H}$ and an internal field $\boldsymbol{H}_M$

$$\boldsymbol{B} = \boldsymbol{H}+\boldsymbol{H}_M = \mu \boldsymbol{H}. \quad \ldots [10.6]$$

The quantity μ determines the enhancement of the magnetic field within a magnet and, for a given magnet, is called the magnetic permeability. The internal field is due to the orbital and spin magnetic moments and is connected with the magnetization by the relation

$$\boldsymbol{H}_M = 4\pi \boldsymbol{I}. \quad \ldots [10.7]$$

From [*6*] and [*7*] we have

$$B = \mu H = H + 4\pi I = H(1 + 4\pi\kappa), \qquad \ldots [10.8]$$

from which

$$\mu = 1 + 4\pi\kappa. \qquad \ldots [10.9]$$

The Temperature Dependence of Magnetic Properties

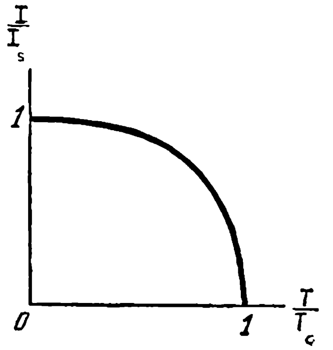


Fig. 175. Temperature dependence of I_s for ferromagnetics.

The diamagnetism of the majority of substances does not depend on temperature, while paramagnetism, with few exceptions (the alkali metals), is strongly temperature dependent. As Curie showed, the magnetic susceptibility of a paramagnetic is inversely proportional to the absolute temperature

$$\chi_A = \frac{C}{T}, \qquad \ldots [10.10]$$

where C is a constant. Curie's law applies to gases and very dilute solutions containing ions of the rare earth elements, that is to those cases where the molecular interaction between the magnetic carriers is weak. In condensed systems when the interaction between magnetic carriers cannot be neglected, Curie's law is replaced by the more general Curie-Weiss law

$$\chi_A = \frac{C}{T-\theta}. \qquad \ldots [10.11]$$

The quantity θ is a constant for a given paramagnetic and is determined by the interaction between magnetic carriers.

In ferromagnetics, the spontaneous magnetization I falls with increasing temperature (Fig. 175) following a curve analogous to that of the temperature dependence of spontaneous polarization of ferroelectrics (Figs. 129*f* and *h*. Compare also Fig. 210). This analogy is not just chance; in this, as in the other cases, one and the same phenomenon is taking place: an order-disorder transition.

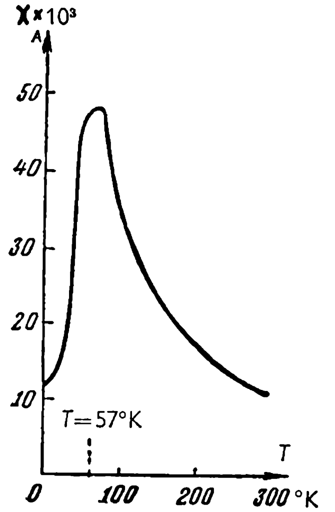


Fig. 176. Temperature dependence of χ_A in iron carbonate.

Antiferromagnetics

The temperature dependence of the magnetic susceptibility of iron carbonate is shown in Fig. 176. At sufficiently high temperatures this substance behaves as an ordinary paramagnetic. As the temperature is lowered χ_A for iron carbonate first increases according to the Curie-Weiss law and then begins to fall. The temperature of maximum susceptibility T_{af} indicates a magnetic transition which, as in the case of the ferromagnetic transition, must be an ordering in the arrangement of the elementary magnets. However, unlike ferromagnetics where the magnets are aligning themselves parallel, in iron carbonate and similar substances an antiparallel orientation is taking place. This leads to cancelling of the magnetic moments and a lowering of the general magnetization of the body. Substances in which antiparallel orientation of the magnetic moment occurs spontaneously are called antiferromagnetics. The possibility of the existence of the antiferromagnetic state was first pointed out by Landau (1933). Ferro- and antiferromagnetism in magnetism are analogous to ferro- and antiferroelectrics in electricity.

General Remarks on Magnets

Diamagnetism is a property of substances, atoms and molecules which do not have magnetic moments of their own so that their magnetization is induced by an external magnetic field. In atoms of diamagnetics, the electron shells are closed and the resultant orbital moments $\boldsymbol{L} = 0$. Usually in these cases, the resultant spin moments are also zero; $\boldsymbol{S} = 0$. Thus the total moment

$$\boldsymbol{I} = \boldsymbol{L}+\boldsymbol{S} = 0. \qquad \ldots [10.12]$$

We can distinguish three groups of diamagnetics.

1. Normal: inert gases, metals of the B sub-group (Cu, Ag, Au, Zn, Cd); the magnetic susceptibility of all these does not depend on temperature and is small in magnitude: $\chi_A \sim 10^{-6}$.

2. Anomalous: bismuth, antimony, gallium, graphite, γ-phases; the magnetic susceptibility here is numerically larger. In graphite: $\chi_{\parallel} = -21{\cdot}5\times 10^{-6}$, $\chi_{\perp} = -0{\cdot}5\times 10^{-6}$. In bismuth: $\chi_{\parallel} = -219\times 10^{-6}$, $\chi_{\perp} = -310\times 10^{-6}$. In crystals with a layered structure, the large value of the magnetic susceptibility is for directions lying in the plane of the layers. This can be readily seen by taking graphite as an example. In bismuth at low temperatures there is a periodic variation of susceptibility with field strength (the de Haas-van Alphen effect) (Fig. 177).

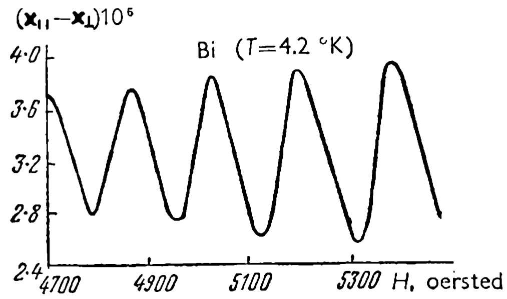


FIG. 177

Dependence of $(\chi_{\parallel}-\chi_{\perp})$ of a single crystal of bismuth on the strength of the magnetic field (for an angle of orientation of the specimen $\phi = 46° 15'$.)

3. Superconductors, in which the magnetic induction $B = 0$. Here we have from [*8*] that $I = -H/4\pi$, whence $\chi = -1/4\pi$. It follows that superconductors can be described as superdiamagnetics. However this description is purely formal since a macroscopic current flows on the surface of superconductors screening the superconductor from the external field.

Paramagnetism is a property of substances, atoms and molecules which do not have fully compensated magnetic moments. The NO molecule, for example, is paramagnetic because it belongs to the rare group of molecules with an odd number of electrons and a consequent uncompensated spin. Molecules with an even number of electrons are diamagnetic, with the exception of the molecules of oxygen and sulphur where paramagnetism is a result of non-compensation of the spins of two electrons (see Chapter 4 § 1). Metals with unfilled *d* and *f* electron sub-groups show large paramagnetism, and the compounds of these metals are also paramagnetic.

Magnetic Properties of the Elements

Table VII (Appendix I) gives values of the magnetic susceptibilities of the elements at room temperature. With few exceptions elements of the A sub-group are paramagnetic and elements of the B sub-group are diamagnetic. Amongst the elements of the A sub-group, the alkali and alkaline earth metals have temperature independent paramagnetism which is much smaller than that of the transition and rare earth metals. Now the cores of the alkali and alkaline earth metals have closed shells so their paramagnetism is due to the free electrons. If the electron gas obeyed classical statistics, then it would be strongly paramagnetic because of the electron spin. Pairing of electrons on the energy levels of the conduction band with consequent mutual cancelling of spins gives diamagnetism. The weak paramagnetism of the electron gas arises from thermal excitation of a small number of electrons which occupy the highest levels of the conduction band.

The ferromagnetics are generally either transition metals belonging to the ferromagnetic triad or rare earth metals—gadolinium and dysprosium—as well as compounds of ferromagnetic elements. However there are also the specially interesting ferromagnetic compounds of non-ferromagnetic components, for example the Heusler alloys $MnCu_2Al$ which have the β-phase structure, and also some other alloys of manganese and of chromium. All these contain elements—manganese, chromium and copper—which lie to left and right of the ferromagnetic triad, and one can say that they are ferromagnetic because their electronic configuration is an approximation to that in a typical ferromagnetic. We note also that the structure of the β-phase which is based on the body-centred cubic lattice is close to that of α-iron. The alloys MnCu and MnAu are antiferromagnetic.

2. Diamagnetism

The Orbital Magnetic Moment

Consider an electron moving in a curved orbit of radius r and describing an area S.

Let ν be the frequency, that is the number of revolutions completed by the electron in 1 second. The orbital magnetic moment is

$$M = \frac{iS}{c} = \frac{e\nu}{c} S, \qquad \ldots [10.13]$$

where c is the velocity of light. The angular velocity ω and linear velocity v of the electron are

$$v = \omega r; \quad \omega = 2\pi\nu. \qquad \ldots [10.14]$$

Putting these values in [*13*] we have

$$M = (e\omega/2\pi c)\,\pi r^2 = (e\omega/2c)\,r = (e/2\mu c)\,\mu v r = (e/2\mu c)\,P, \qquad \ldots [10.15]$$

where P is the orbital mechanical moment and μ is the mass of the electron. Quantization of the mechanical moment

$$P = (h/2\pi)\,m \qquad \ldots [10.16]$$

leads to quantization of the orbital magnetic moment

$$M = \frac{eh}{4\pi\mu c}\,m = mM_B, \qquad \ldots [10.17]$$

where

$$M_B = eh/4\pi\mu c \qquad \ldots [10.18]$$

is the Bohr magneton (the quantum of magnetic moment).

From formula [*15*] we can derive the gyromagnetic ratio of the orbital moment

$$M/P = e/2\mu c. \qquad \ldots [10.19]$$

This ratio is twice the ratio of the spin moments [*2.2*]. This difference between the orbital and spin gyromagnetic ratios is the main clue to the origin of ferromagnetism.

The magnitude of the orbital moment depends on angular velocity and it follows from [*15*] that a change of $\Delta\omega$ in angular velocity gives a change in magnetic moment

$$\Delta M = (er^2/2c)\,\Delta\omega \qquad \ldots [10.20]$$

Universality of Diamagnetism

Diamagnetism is a universal property of all bodies, but since diamagnetic effects are so small, they can be observed only in those cases when there are no other, stronger magnetic effects. This specially applies to paramagnetics where diamagnetism is concealed by paramagnetism which is much greater in absolute magnitude. Diamagnetism is connected with the orbital motion of electrons in atoms as well as in molecules and crystals, and is due to small perturbations in the angular velocity of orbital motion which occur when the atom is brought into a magnetic field. What happens can be seen from the following general considerations: the motion of an electron in orbit can be thought of as a closed current loop. When this loop is brought into a magnetic field then Faraday's law of electromagnetic induction tells us that additional e.m.f. is developed in the loop and that the current in the loop is thereby changed, producing an additional magnetic moment. According to Le Chatelier's principle this magnetic moment is of such a direction as to oppose the applied field which produces the induction. This shows that the induced moment is directed against the field and this then is the basic indication of a diamagnetic. From Le Chatelier's principle also, it follows that a diamagnetic must move out from a magnetic field.

Frequency of Precession of an Orbit in a Magnetic Field

An electron moving in a curved orbit experiences a centripetal force $f = \mu v^2/r$. When a magnetic field H is applied perpendicular to the plane of the orbit (Fig. 178*a*), an additional Lorentz force $\Delta f = -evH/c$ arises parallel to the radius. The resulting centripetal force is then

$$f_n = f + \Delta f = \mu v_1^2/r = \mu v^2/r - evH/c,$$

from which we get

$$\mu(v_1^2 - v^2)/r = \frac{\mu(v_1+v)(v_1-v)}{r} \approx \frac{2\mu v}{r}\Delta v = -\frac{evH}{c}.$$

The change in linear velocity of orbital motion by Δv corresponds to a change of angular velocity by

$$\Delta\omega = \Delta v/r = -eH/2\mu c. \qquad \ldots [\mathit{10.21}]$$

In the general case, if the orbit is inclined to the field (Fig. 178*b*) then the magnetic

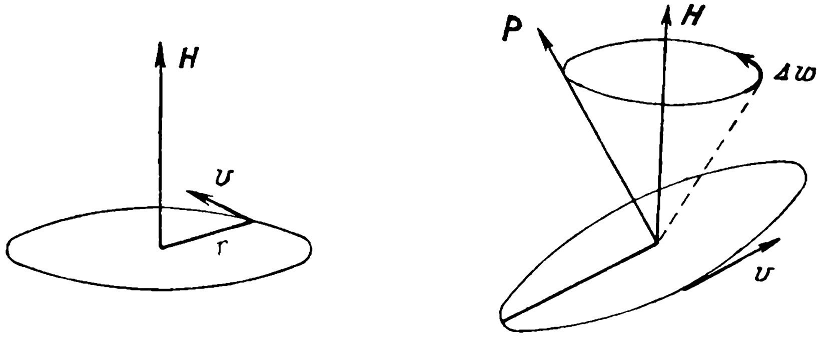


FIG. 178. Calculation of the angular velocity of precession of an electron orbit.

field causes the orbit to precess: the normal $\boldsymbol{P}$ to the plane of the orbit describes a cone round the magnetic field vector $\boldsymbol{H}$. The angular velocity of precession is $\Delta\omega$ as given by formula [*21*].

Diamagnetism of Atoms

Precessional motion of an orbit in a magnetic field has the effect of producing an additional magnetic moment the magnitude of which is obtained from formula [*20*] by substituting for $\Delta\omega$ from [*21*]

$$\Delta M = (eS/2\pi c)\,\Delta\omega = -(e^2 S/4\pi\mu c^2)\,\boldsymbol{H}. \qquad \ldots [\mathit{10.22}]$$

The negative sign shows that the induced magnetic moment is in a direction opposed to the field. For closed shells we may assume spherical symmetry and consider circular orbits, for which $S = \pi r^2$. For a spherically symmetric atom $\overline{r^2} = \frac{3}{2}\, r^2$. Applying formula [*22*] to all the electrons of the atom we get the total magnetic moment of the atom

$$\Delta M = -(e^2 H/6\mu c^2)\sum_{k=1}^{Z} r_k^2, \qquad \ldots [\mathit{10.23}]$$

where Z is the number of electrons. Hence for the molar magnetic susceptibility we have, remembering that $r_k^2 \simeq 10^{-16}\,\mathrm{cm}^2$

$$\chi_A = (N/M)\Delta H = -(Ne^2/6\mu c^2)\sum_{k=1}^{Z} r_k^2 = Z \times 10^{-6}. \qquad \ldots [\mathit{10{\cdot}24}]$$

The order of magnitude of the diamagnetic susceptibility determined from [*24*] agrees with experiment. Diamagnetism of atoms does not depend on temperature.

Diamagnetism of the Electron Gas

The magnetic moment of free electrons in a metal associated with their motion is zero since the electrons move in straight lines between collisions and the moment of the total motion is zero. An external magnetic field changes the character of the electron motion, converting the straight line orbits into curved ones. As a result, a magnetic moment analogous to the orbital moment becomes associated with each electron in the electron gas. However, if we calculate the total effect of all the electrons, it turns out that the induced diamagnetic moment within the metal is completely cancelled out by a moment in the reverse direction due to electrons reflected from the surface of the metal. The total effect of the two currents (*a*) that induced by the magnetic field in the electron gas and (*b*) that of the reflected electrons can be taken to be that of two peripheral currents flowing against each other (Fig. 179). This can also be obtained from energy considerations: the Lorentz force, perpendicular to the velocity of motion of the electron, changes only the shape of the trajectories and does not change their energy. Thus the application of a magnetic field does not change the energy of the classical electron gas, so diamagnetism of the electron gas in this case must be zero. Landau (1930) showed that an essentially different situation obtains in quantum theory because of quantization of periodic motion and that quantization changes the energy of the electron gas for the diamagnetic susceptibility of which Landau obtained the formula

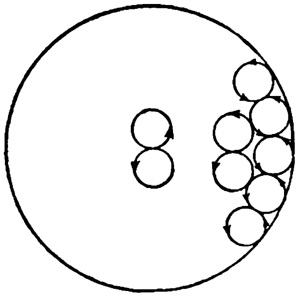


FIG. 179
Mutual cancellation of the diamagnetism of free electrons.

$$\chi_A = -\tfrac{1}{3} N' M_B^2 / kT, \qquad \dots [10.25]$$

where N' is the number of conduction electrons contributing to diamagnetism. Since most levels in the conduction band are full, only a small fraction of the electrons can change their energy, i.e. those electrons which lie in a narrow region of thermal activation near to the upper limit of the full zone, so

$$N' = N (kT/\varepsilon_{\max}), \qquad \dots [10.26]$$

where $\varepsilon_{\max}$ is the maximum kinetic energy of the electrons and does not depend on temperature.

From [*25*] and [*26*] we obtain

$$\chi_A = -\tfrac{1}{3} N M_B^2 / \varepsilon_{\max}. \qquad \dots [10.27]$$

Introducing the value of $\varepsilon_{\max}$ [*8.17*] we obtain

$$\chi_A = -\frac{4\mu M_B^2}{h^2}\left(\frac{\pi}{3}\right)^{2/3} N^{1/3}. \qquad \ldots [10.28]$$

Formula [*28*] shows that the free electron gas shows temperature independent diamagnetism, the magnitude of which is very small.

Diamagnetism of Metals

Diamagnetism of metals is obtained from the diamagnetism of their atomic cores, by use of formula [*24*] and from the diamagnetism of their free electrons (formula [*28*]). We may thus expect that the diamagnetism of metals will be greater than the combined diamagnetism of the metal ions. The experimental data are given below.

Diamagnetism of Metals and Metallic Ions

Metal	χ_A	Ion	χ_A
Cu	$-5{\cdot}5\times10^{-6}$	Cu^+	$-18{\cdot}0\times10^{-6}$
Ag	$-21{\cdot}6$	Ag^+	$-31{\cdot}0$
Au	$-29{\cdot}6$	Au^+	$-45{\cdot}8$

These data show that diamagnetism of metals is less than diamagnetism of their ions. Thus it follows that the electron gas must be paramagnetic as a result of the constant spin moments of the electrons. That diamagnetism of metals is independent of temperature is evidence for the temperature independence of the paramagnetism of the electron gas.

3. Paramagnetism

Langevin's Theory of the Temperature Dependence of Paramagnetism

Consider a gas of paramagnetic molecules with constant magnetic moment μ_0. The energy of magnetic interaction of two elementary magnets $\sim 10^{-16}$ erg, while thermal energy at $T = 10°$ K is $kT \sim 10^{-15}$ erg. Thus in the absence of an external field, thermal motion disturbs the orientation of magnetic moments and the resultant magnetization is $I = 0$. When a magnetic field is applied the magnetic moments become orientated parallel to the field. The magnetic energy of a magnet in an external field (neglecting the interaction of the elementary magnets with each other) is

$$U = -\mu_0 H \cos\theta. \qquad \ldots [10.29]$$

This energy has a minimum value when $\theta = 0$, where θ is the angle between the direction of the magnetic moment and the direction of the field. The degree of orientation of the moments and the magnitude of the resultant moment is determined by the ratio of the magnetic energy $\mu_0 H$ to the thermal energy kT. The probability that an elementary magnet will be aligned at an angle θ is determined by the Boltzmann function

$$w = \exp(-U/kT) = \exp[(\mu_0 H/kT)\cos\theta] = \exp(\beta\cos\theta) \qquad (\beta = \mu_0 H/kT) \qquad \ldots [10.30]$$

The mean magnetic moment

$$\mu = \overline{\mu_0 \cos\theta} = \mu_0 \frac{\int\limits_0^\pi \cos\theta e^{\beta\cos\theta}\sin\theta d\theta}{\int\limits_0^\pi e^{\beta\cos\theta}\sin\theta d\theta} = \mu_0 \frac{d}{d\beta}\ln\int\limits_0^\pi e^{\beta\cos\theta}\sin\theta d\theta =$$

$$= \mu_0 \frac{d}{d\beta}\ln\frac{e^\beta - e^{-\beta}}{\beta} = \mu_0 \frac{\beta}{e^\beta - e^{-\beta}}\left\{\frac{e^\beta + e^{-\beta}}{\beta} - \frac{e^\beta - e^{-\beta}}{\beta^2}\right\} = \mu_0\left(\frac{e^\beta + e^{-\beta}}{e^\beta - e^{-\beta}} - \frac{1}{\beta}\right).$$

Whence

$$\bar{\mu} = \mu_0(\coth\beta - 1/\beta). \qquad \ldots [10.31]$$

For weak fields and not too high temperatures the parameter $\beta = \mu_0 H/kT \ll 1$ and so [*31*] can be expanded in series giving

$$\bar{\mu} = \mu_0(\tfrac{1}{3}\beta + \ldots) = \mu_0^2 H/3\,kT. \qquad \ldots [10.32]$$

Hence for a gram molecule of the substance we have

$$\chi_A = N\bar{\mu}/H = N\mu_0^2/3\,kT. \qquad \ldots [10.33]$$

This is Curie's law. Formula [*33*] enables us to determine the constant C in [*10*] in terms of molecular constants. Introduction of the laws of spatial quantization into the classical Langevin theory takes account of the discrete orientation of the dipoles, but does not change the final result.

The Weiss Theory

Weiss suggested that the external field H which causes the magnetization of a body results in the production of an ordered internal magnetic field proportional to the magnetization [*7*]. Thus within a ferromagnetic there exists an effective field

$$H_{\text{eff}} = H + cI.$$

Introducing the formula for H_{eff} into the Langevin formula we get

$$I = N\mu_0^2 H_{\text{eff}}/(3\,kT) = N\mu_0^2 H/(3\,kT) + Nc\mu_0^2 I/(3\,kT),$$

whence

$$\chi_A = \frac{I}{H} = \frac{N\mu_0^2}{3k\left(T - \frac{Nc\mu_0^2}{3k}\right)} = \frac{C}{T - T_k}. \qquad \ldots [10.34]$$

This is the Curie-Weiss law.

Paramagnetism of the Electron Gas

Since electrons have an intrinsic magnetic moment, the electron gas is paramagnetic. In a magnetic field the electron spins become oriented preferentially in the direction of the field, thus giving a resultant magnetic moment. The classical Langevin formula does not serve to calculate the paramagnetism of the electron gas since the paramagnetism of a variety of metals does not depend on temperature. To calculate the paramagnetism of the free electrons we have to use Fermi statistics. Thus all levels in the conduction band will be filled in a zone of width $\varepsilon_{\max}$ where $\varepsilon_{\max}$ is the maximum kinetic energy of the electrons; on each level there will be two electrons with opposite directions of spin. Let us divide the occupied band into two half bands thus separating the spin directions of the electrons (Fig. 180*a*). In a magnetic field the level of the half-band with parallel spin is lowered by $-M_BH$ and the level of the half-band with antiparallel spin is raised by $+M_BH$ (Fig. 180*b*). The difference in energies of the half bands is $2M_BH$. The difference between the levels to which the half bands are filled leads to a flow of electrons and equalization of the level (Fig. 180*b*). As a result there is an increase in the number of electrons with parallel spins and a paramagnetic moment arises, which is

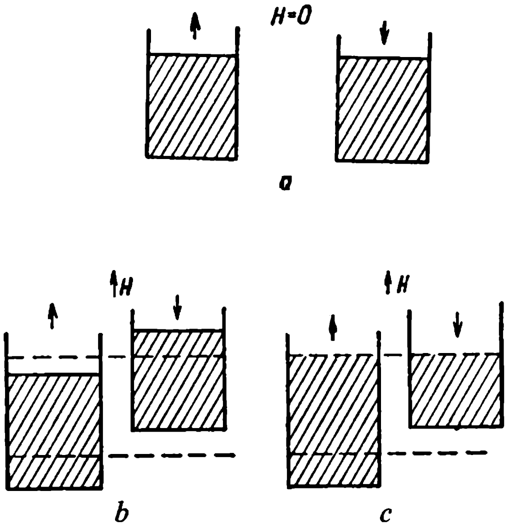


FIG. 180
Calculation of the paramagnetism of free electrons.

$$M = \Delta Z M_B, \qquad \ldots [10.35]$$

where ΔZ is half of the difference between the initial numbers of levels of the half bands. The deficiency in the kinetic energy of the electrons in the left half-zone is equal to the gain in magnetic energy

$$\Delta Z \,.\, \varepsilon_{\max} = M_BH. \qquad \ldots [10.36]$$

Combining [*35*] and [*36*] we obtain the magnetic susceptibility

$$\chi_A = NM/\varepsilon_{\max} = NM_B^2/\varepsilon_{\max}. \qquad \ldots [10.37]$$

Comparison of formulae [*27*] and [*37*] shows that paramagnetism of the electron gas is three times larger than its diamagnetism and the difference is the paramagnetism of the conduction electrons. In this way we can explain the weak paramagnetism of the alkali metals.

Paramagnetism of the Transition Metals

The large paramagnetism of the transition and rare earth metals is connected with their *d*- and *f*-subgroups. In accordance with Hund's rules the electrons fill these subgroups in such a way that the total spin moment is a maximum. If the *d*- or *f*-subgroups are not full this leads to a resultant magnetic moment for the atom and to this an orbital moment also has to be added. In the transition metal group, as is evident from Fig. 181, there is a steep rise in paramagnetic susceptibility as we approach the iron triad. In the ferromagnetic metals iron, cobalt and nickel at sufficiently low temperatures ($T \to 0$) there is a saturation magnetization, I_{sat} even in weak fields. Dividing I_{sat} by the number of atoms we obtain the magnetic moment of the atoms

Metal	Fe	Co	Ni	
atomic magnetic moment	2·2	1·7	0·6 M_B.	. . . [*10.38*]

As the atomic number of the metal in the ferromagnetic triad is increased the magnetic moment becomes less. This is connected with the filling of the *d*-subgroup. In the following element, copper, the *d*-subgroup is completely filled, which explains its diamagnetic properties. The number of valence electrons of the nickel atom is sufficient to fill the *d*-band but the presence of a paramagnetic moment shows that, on the average, 0·6 electrons per atom in metallic nickel have gone into the *s*-band. This is confirmed by the magnetic properties of alloys of paramagnetic nickel with diamagnetic metals and with hydrogen.

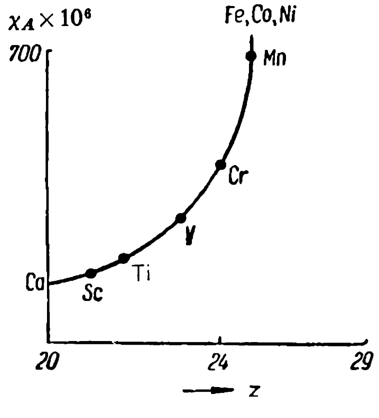


FIG. 181
Magnetic susceptibility of transition metals of the fourth period.

Fig. 182*a* shows the changes in magnetic susceptibility of palladium/hydrogen alloys. With increasing concentration of hydrogen, paramagnetism is reduced approximately linearly and vanishes at 60 at. % of hydrogen. This can be explained by the filling of the *d*-band of the alloy by the valence electrons of the atoms of hydrogen, remembering that the alloy is a solid solution of hydrogen in palladium. In Fig. 182*b* similar results are given for the alloys of nickel with copper and with zinc. Magnetic properties of the alloys of nickel with monovalent copper are similar to those of the alloys of palladium with hydrogen. In the alloys of nickel with divalent zinc the extinction of the paramagnetic moment of the nickel atom takes place twice as fast and the paramagnetism vanishes at 30 at. % of zinc, that is at half the concentration that is found in alloys with monovalent elements. The similarity of the results for the alloys of nickel and those for its homologue palladium shows that the *s*-bands in both of these are filled to the same level.

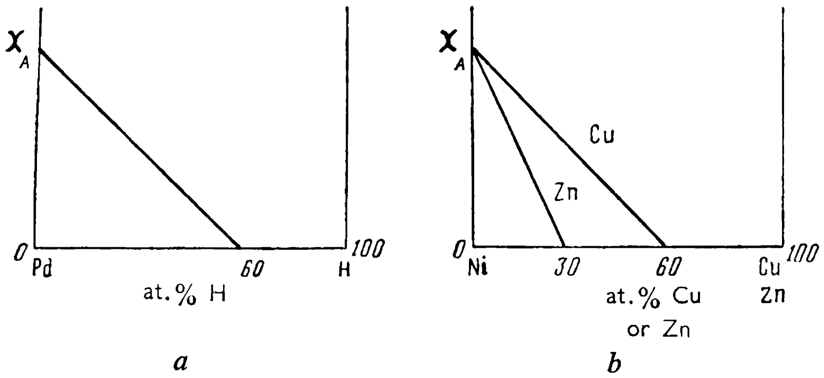


FIG. 182. Paramagnetic susceptibility of alloys: (*a*) palladium based, (*b*) nickel based.

Application of Paramagnetism to obtain Low Temperatures

When a paramagnetic is magnetized adiabatically its temperature is increased and the entropy due to ordering falls. Demagnetization produces a disordering of the magnetic structure and lowering of the temperature by an amount which may be large since the thermal capacity of solids at low temperatures is small. Using a solid solution of chromium potassium and aluminium potassium alums K_2SO_4 (Cr, Al) $(SO_4)_3 . 24H_2O$ with a volume of 56 cm^3, demagnetization at $T = 1{\cdot}174°$ K from a field of 24,000 Oe. to one of 1 Oe., gave a temperature of $T = 0{\cdot}0044°$ K. Evaporation of liquid helium gives a much higher temperature: $T = 0{\cdot}71°$ K.

4. Ferromagnetism

The Gyromagnetic Anomaly

The Einstein-de Haas experiment (1915) showed that magnetization of a ferromagnetic needle was accompanied by the development of a torsional moment. The gyromagnetic effect as demonstrated by this experiment shows the essential nature of magnetic gyroscopes in a ferromagnetic. Changes in the magnetization of the needle were connected with reversals of the direction of the axes of the gyroscopes and these are accompanied by changes in the total mechanical moment of the body. If the direction of magnetization was changed periodically with the frequency of free torsional oscillations of the needle then torsional oscillations with measurable amplitude were set up. The converse experiment, that of producing magnetization by twisting a ferromagnetic was performed by Barnett (1915). From gyromagnetic experiments such as these we can determine the gyromagnetic ratio M/P. Barnett made measurements with various ferromagnetic materials (iron, cobalt, nickel and the Heusler alloys) and showed that magnetization was connected with negative charges, that is with electrons. The magnitude of the gyromagnetic ratio was shown to be

$$\frac{M}{P} = \frac{e}{\mu c},$$

that is, twice as large as the gyromagnetic ratio to be expected for orbital moments [*19*]. Until the discovery of the spin of the electron (1926) the reason for this divergence was unknown and it was described as an anomaly. But since discovery of the magnetic moment of the electron and the half-integral magnitude, $s = \frac{1}{2}$, of its spin quantum number it is clear that magnetization of a ferromagnetic is connected with electron spins and not with orbital moments.

The Regions of Spontaneous Magnetization

A characteristic of a ferromagnetic is the large magnetization which reaches saturation even in weak magnetic fields. This saturation magnetization I_s depends on temperature and vanishes at the Curie point. The curve of $I_s = I_s(T)$ is similar to the curve of atomic disordering and to the curve of spontaneous polarization of a ferroelectric. In all these cases there exist regions (domains) of spontaneous order, either of atomic structure or of electric or magnetic moments. The presence of domains in a ferromagnetic is readily demonstrated by means of magnetic suspensions (Bitter patterns). Within the limits of each domain the direction of spontaneous magnetization is uniform. In neighbouring domains the directions are different. Domains break down as a result of thermal motion and grow when the temperature is reduced. The boundaries of the domains are not in the same places as the boundaries of crystals or of crystal mosaic blocks. During magnetization, reversals occur in the orientation of the domain axes. To study the stages in the process of transformation one can observe the reversal of individual domains. This occurs by discontinuous changes of the magnetization (Barkhausen effect).

The Magnetization Curve

Because of anisotropy, the magnetization for a given field strength depends on the orientation of the crystal relative to the direction of the field (Fig. 183*a*). In the absence of a field, the axes of the domains in the individual crystals are aligned along the directions of easy magnetization. In a nickel crystal this is the [111] direction.

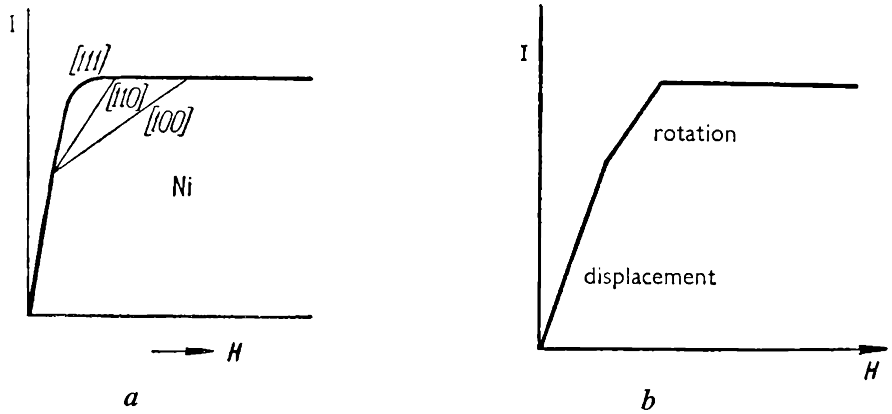
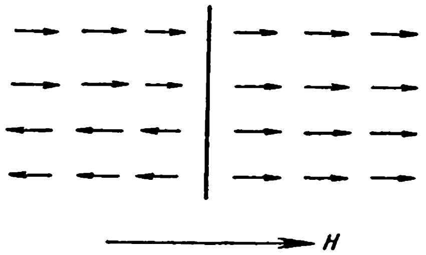


FIG. 183. The magnetisation of (*a*) a nickel single crystal, (*b*) polycrystal.

FIG. 184
Reversal of domain directions on magnetization.



After this comes the [110] direction and finally [100]. The magnetization curve of a polycrystal is the sum of the magnetization curves of the individual crystals and in the absence of texture, that is preferred orientation of the crystallites, the effects of magnetic anisotropy are smoothed out. Nevertheless one can distinguish sections on the magnetization curve of a polycrystal corresponding to the different stages of magnetization (Fig. 183*b*). In weak fields displacements take place, that is growth of domains with favourable orientation at the expense of neighbouring regions with unfavourable orientation (Fig. 184). As the field is increased rotation takes place, that is there is a tendency to align the magnetic moments in the direction of the field and not in general in the direction of easy magnetization. In fields larger than the practical saturation value, when all the moments of the domains are parallel to the field, orientation of individual electron spins still take place. This is obstructed by thermal motion. This process is called true magnetization.

The Nature of Ferromagnetism

If we start from the supposition that the internal molecular field H_M is the cause of spontaneous magnetization then we must expect that this field will be very large. However phenomena observed (Ya. G. Dorfman) when β-particles pass through a thin foil of a ferromagnetic do not give support to this supposition. It turns out that the forces which orient the electron spins parallel to each other are not magnetic in origin but are forces of electrical interaction in a system of electrons obeying quantum laws. The type of orientation of the electron spins is determined by the sign of the exchange integral A which can be calculated by the method of atomic orbitals (Chapter 3 §3) and for a system of two electrons is

$$A = e^2 \int \frac{\psi_n^*(r_1)\psi_m(r_1)\psi_n(r_2)\psi_m^*(r_2)}{r_{12}}\,dv_1 dv_2, \qquad \ldots [10.39]$$

where ψ_n and ψ_m are unperturbed electron wave-functions and r_1 and r_2 are radius vectors of the two electrons. If the exchange integral is negative then there is antiparallel orientation of spins; this usually is found in covalent binding (the water molecule) and in antiferromagnetics. To a positive integral corresponds parallel orientation of spins; this is found during the filling of the electron subgroups in atoms (compare Hund's rule for maximum spin moment) and in ferromagnetics. The sign of the exchange integral depends on the distance between the electrons. In metals the integral is negative for small separation and positive for larger separations. Fig. 185 shows the magnitude of the exchange integral as a function of interatomic distance. $A > 0$ for the ferromagnetic modification of iron, for cobalt and for nickel. The magnitude of the exchange integral is connected with the Curie temperature in that the larger is A, the greater is the thermal energy kT_c necessary to destroy ferromagnetism. The relative positions on the curve of $A = A(r)$ of the points for the metals of the ferromagnetic triad correspond to the variation of their Curie temperatures:

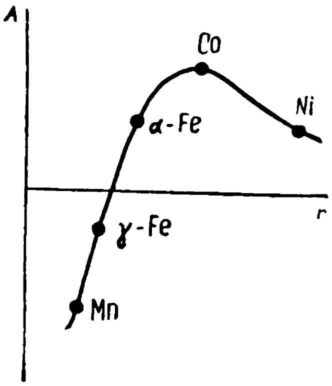


FIG. 185
Curve of the dependence of exchange energy on separation.

Metal	Fe	Co	Ni	Gd
T_c	768	1150	365	16 °C

Parallel orientation of the spins in the range of spontaneous magnetization forces the electrons to occupy higher levels in the energy band, thus increasing their kinetic energy.

The above explanations allow us to formulate the conditions favourable for the occurrence of ferromagnetism. 1. The electron shells in the atoms must not be completely filled. 2. The density of states in these shells must be large so that the increase in kinetic energy produced by occupying higher vacant levels does not exceed the reduction due to the exchange energy.

Electrons from structures with full shells do not satisfy the first requirement and electrons of the non-binding valence shells in which the density of states is small, do not satisfy the second requirement. In the majority of metals, the conditions required for the occurrence of ferromagnetism are not fulfilled. Thus we comparatively rarely find traces of ferromagnetism in pure metals, the most favourable cases being found among the transition metals and lanthanides which have unfilled internal *d*- and *f*-subgroups. The occurrence of ferromagnetism requires optimum interatomic spacing: if the atoms are too far apart then the exchange interaction is very small; if the atoms are too close together the energy band of the electron subgroup is broadened, giving a rapid rise in kinetic energy. This is shown by the ratio of atomic radius to the radius of the *d*-shell.

Metal	Ti	Cr	Mn	Fe	Co	Ni
r_a/r_d	1·12	1·18	1·47	1·63	1·82	1·98

For transition metals of the fourth period the optimum value of this ratio is somewhat greater than 1·5.

Magnetic interaction, as was shown above, is small compared with exchange interaction. Nonetheless the magnetic interaction gives rise to a series of important phenomena: the magnetic anisotropy, magnetostriction, hysteresis and the technical processes of magnetization.

Magnetic Moments of Atoms of the Transition Metals

The maximum spin moment in the *d*-subgroup is 5 and corresponds to the filling of one half-band. Let n be the total number of electrons in the *s*- and *d*-bands, of this x electrons are in the *s*-band and $(n-x)$ electrons in the *d*-band. Suppose that in one half-band there are 5 electrons and in the other half-band $n-x-5$. The resultant spin moment of the electrons in the *d*-subgroup will be

$$M_d = 5-(n-x-5) = 10+x-n. \qquad \ldots [10.40]$$

Assuming a uniform degree of filling of the *s*-band for the various transition metals and taking $x = 0{\cdot}6$, we obtain the following theoretical values of the atomic spin moments:

metal	Ni	Co	Fe
n	10	9	8
$M_d = 10{\cdot}6-n$	0·6	1·6	2·6

. . . [*10.41*]

This agrees fairly well with experimental data (see [*38*] above).

In Fig. 186 the line $M = 10{\cdot}6-n$ is shown. The part of the line AB reproduces well the magnetic moments of iron, cobalt, nickel and their alloys with copper. The part BC does not agree with experiment. In the region of the point B for which $n = 8{\cdot}2$ and $M_d = 2{\cdot}4$ the atomic moment reaches a maximum and then falls along the curve BD. This shows that one half-band cannot be completely empty when the second half-band is nearly full. When the degree of filling of the 'empty' half-band is reduced, about 2·6 electrons must remain in it. Fig. 187 shows, in accordance with this explanation, the electron distribution in the *d*-band of transition metals, from which it follows that the magnetic moment of manganese is 1·2, and that of chromium 0·2.

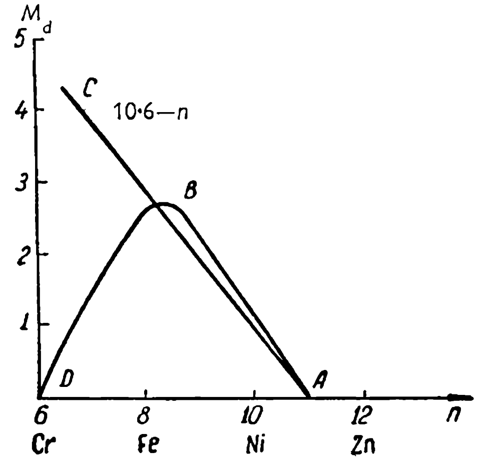
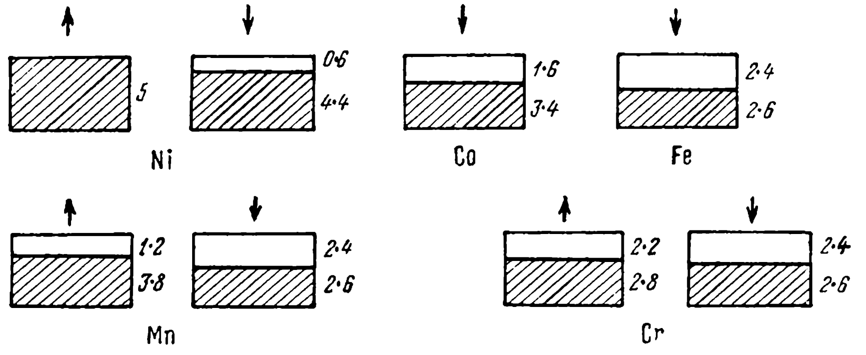


Fig. 186
Magnetic moments of atoms of transition metals.

Fig. 187
Filling of the *d*-bands in transition metals.



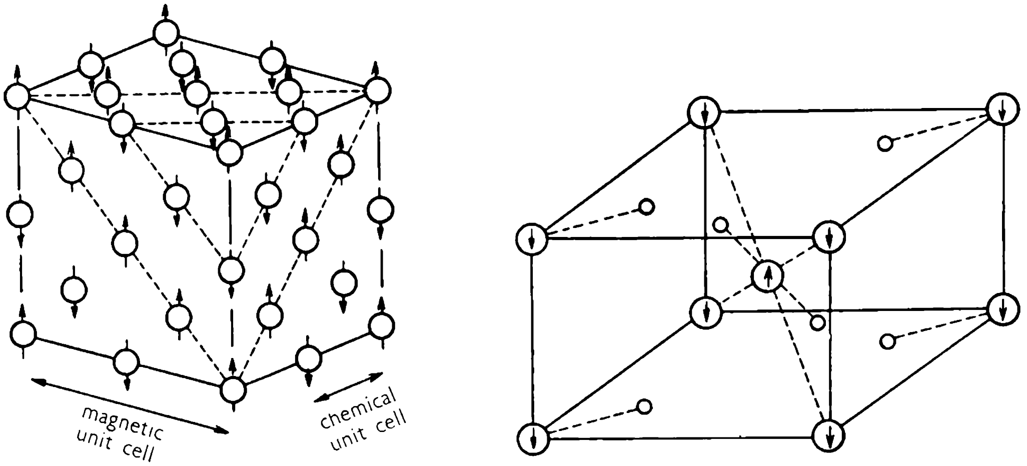
5. Antiferromagnetism and Magnetic Structure

Magnetic Structure and its Connection with Atomic Structure

At the basis of ferro- and antiferromagnetism there is one underlying phenomenon—ordering in the arrangement of magnetic moments. In an antiferromagnetic the antiparallel arrangement is more favourable. Here magnetic dipoles are connected with the lattice points of the atomic structure, so where antiferromagnetic structure occurs we can speak of the magnetic structure of the body. It turns out that the magnetic structure can be determined by means of neutron diffraction. In Fig. 188 are shown the magnetic structures of two antiferromagnetics: MnO ($T_c = 120°$ K) and MnF_2 ($T_c = 80°$ K) as determined by neutron spectroscopy at low temperatures.

Manganese oxide has a structure of the NaCl type. Fig. 188 shows only the arrangement of the manganese ions which form a face-centred structure. The magnetic structure which occurs at low temperatures can be thought of as a succession of planes belonging to one octahedral $\{111\}$ system alternately with parallel and antiparallel orientation of the moments along one of the cube edges. As a result of this, the period of the magnetic structure in some directions is twice the period of the atomic structure. The magnetic structure of MnO can be considered as a superlattice very similar to the superlattice which occurs in the alloy CuPt (see Tables 33 and 34). The magnetic interaction of the octahedral planes with different orientations of their magnetic moments leads to a small rhombohedral compression of the atomic structure of MnO which can be observed by X-rays.

Fig. 188*b* shows the magnetic structure of manganese difluoride; the large circles with arrows in them represent ions of manganese, the small circles, ions of fluorine. The manganese ions at the vertices and in the centre of the tetragonal cell have antiparallel orientations with dipole direction along the tetragonal axis. As a result of this there is a change of translational group in that the period along the body diagonal of the unit cell of the magnetic structure is doubled as compared with the period of the atomic structure. This leads to a change in the type of Bravais Lattice which is transformed from body-centred to simple.



a *b*

FIG. 188. Magnetic structures of antiferromagnetics below the Curie temperature: (*a*) MnO, (*b*) MnF_2.

Origin of Ferromagnetism of Ferrites

In conclusion we must stress the close relation between ferro- and antiferromagnetism, and in particular the possibility of the occurrence of ferromagnetism as a result of incomplete compensation of antiferromagnetism, such as takes place in ferrites. Ferrites are double metal oxides with the composition $MO \,.\, Fe_2O_3$, where M = Mg, Zn, Cu^{++}, Ni^{++}, Co^{++}, Fe^{++}, Mn^{++}—all divalent metals. All have the cubic structure of the spinel ($MgAl_2O_4$) type (Fig. 189) belonging to the space group O_h^7-$Fd3m$. The unit cell comprises 8 formula units, that is 32 oxygen atoms, 8 atoms of type M and 16 atoms of trivalent iron. The oxygen atoms occupy positions 32*e* and take up cubic close packing. The cations are situated on the octahedral (8*b*) and tetrahedral (16*c*) spaces of this packing.

Structural investigations have shown that in general there are two types of atomic distribution corresponding to the normal and inverse spinels:

type of spinel structure	8*b*	8*c*
normal	8*M*	$16Fe^{+++}$
inverse	$8Fe^{++}$	$8M+8Fe^{+++}$.

In the inverse spinel, half the iron atoms occupy octahedral sites 8*b* and the tetrahedral positions are filled either by iron atoms or by atoms of the divalent metal. A large number of ferrites have the inverse structure; these include the ferrites of manganese, iron, cobalt and nickel. Zinc and cadmium ferrites have the normal structure. The cations which possess uncompensated magnetic moments are distributed on the

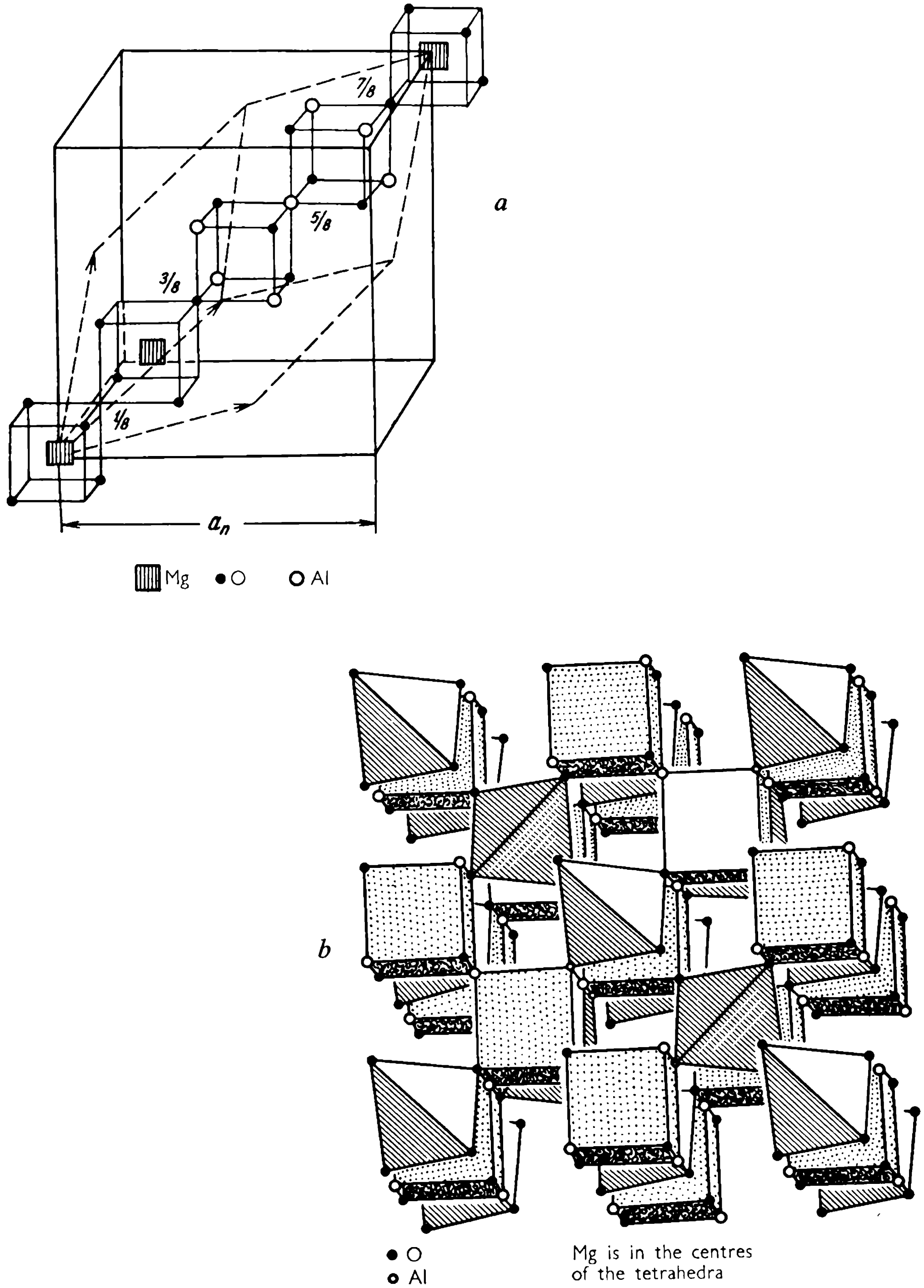


FIG. 189. The spinel-type structure of $MgAl_2O_4$: (*a*) atomic arrangement in the unit rhombohedron, (*b*) atomic arrangement in polyhedral representation.

octahedral and tetragonal sites in antiparallel orientation as shown by the following scheme

$$\begin{array}{lll} 8b & 16c & \\ \rightarrow & \rightarrow & \leftarrow \\ \mathrm{Fe}^{+++} & (\mathrm{M} & \mathrm{Fe}^{+++}). \end{array} \qquad \ldots [10.42]$$

As a result, the magnetic moments of the trivalent iron ions cancel each other out and if the divalent M ion has an intrinsic magnetic moment then this produces the uncompensated spontaneous magnetic moment of the ferrite similar to the spontaneous magnetization of a ferromagnetic.

The theory of ferromagnetism of ferrites proposed by Néel (1948) satisfactorily explains their properties. It follows from the electron configuration of the ions that the divalent metal ions have the following intrinsic magnetic moments:

ion	Mn^{++}	Fe^{++}	Co^{++}	Ni^{++}	Cu^{++}	Zn^{++}
μ_0 in M_B	5	4	3	2	1	0.

The magnetic moment of the trivalent iron ion is $\mu_0 = 5\ M_B$ and the moments of the divalent ions of magnesium and cadmium as well as zinc are zero. Table 62 gives the total spin moments calculated on the above scheme. These moments are in acceptable agreement with experimental results.

Table 62

Magnetic Moments of Ferrites

Ferrite		Orientation of spin moments	Sum of the moments	Experimental value
8*b*	16*c*	8*b* 16*c* $\rightarrow\ \rightarrow\ \leftarrow$		
Fe^{+++}	$(Mn^{++}Fe^{+++})$	5 (5 5)	5	5·0
Fe^{+++}	$(Fe^{++}Fe^{+++})$	5 (4 5)	4	4·2
Fe^{+++}	$(Co^{++}Fe^{+++})$	5 (3 5)	3	3·3
Fe^{+++}	$(Ni^{++}Fe^{+++})$	5 (2 5)	2	2·3
Zn^{++}	$(Fe^{+++}Fe^{+++})$	0 (5 5)	0	0
Cd^{++}	$(Fe^{+++}Fe^{+++})$	0 (5 5)	0	0

Chapter 11

PHASE EQUILIBRIUM AND PHASE TRANSFORMATIONS IN ALLOYS

1. Basic Conceptions

General Conditions for Equilibrium

Consider a system consisting of a mixture of atoms of various kinds—for simplicity a binary system of atoms A and B. In the solid state, various cases of phase equilibrium are possible. It may be that the system is homogeneous (one phase) being either a solid solution (A, B) or a chemical compound A_nB_m. Or the system may be heterogeneous consisting of a mixture of various phases, for example if the components are insoluble in each other it may consist of a mixture of crystals of the pure components A + B or it may consist of a mixture of crystals of terminal solid solutions and so on.

There is a large number of phase states of a system. They can be determined by experiment or by calculation using thermodynamic and statistical relations if we know the chemical composition described by the concentration c, the thermodynamic parameters (temperature T, pressure p, volume v, and if necessary even other parameters such as the surface energy etc.). We also need the basic thermodynamic functions for the substance; its internal energy E and entropy S.

The equilibrium of a mechanical system is determined by the minimum of its potential energy: $U = \min$. Thermal equilibrium is determined by equalising the temperature of all bodies which are in thermal contact: $T = \text{const}$. Chemical equilibrium is determined by the constancy of concentration of all the phases taking part in a chemical reaction: $c_i = \text{const}$. Any changes in the system, mechanical, physical or chemical which show themselves as changes with time of pressure, volume, temperature or concentration are evidence of departure from equilibrium. The general conditions for equilibrium are the absence of such changes.

Internal Energy and Entropy

The internal energy E of a system is the sum of the potential (U) and kinetic (K) energies of the particles of the system.

$$E = U + K = E(p, v, T). \qquad \ldots [11.1]$$

The energy E is a function of the state of the system determined by the parameters p, v, T. These parameters are related to each other by the equation of state

$$f(p, v, T) = 0,$$

and in consequence there are only two independent parameters in equation [*1*]. A change in the internal energy of a system may cause a change in its temperature. In such a system let a quantity of heat δQ be given to it or taken from it. A change in temperature leads to a change in one or both of the other thermodynamic parameters p and v. In changing its volume by δv the system does work of expansion $\delta A = p\delta v$. On the basis of the law of conservation of energy

$$\delta Q = dE + \delta A.$$

Changes in the internal energy depend only on the initial and final states of the system. This requires that dE be a perfect differential whilst the work δA done in the system and the heat δQ given to it or removed from it are not perfect differentials.

Another important thermodynamic function is the entropy S, changes of which are determined by the relation $dS = \delta Q/T$, again a perfect differential. Apart from an additive constant, the entropy is a one-valued function of the state. The magnitude of this constant is determined by the Nernst theorem according to which the entropy of a system is zero when the system is in equilibrium at absolute zero.

The Free Energy of Condensed Systems

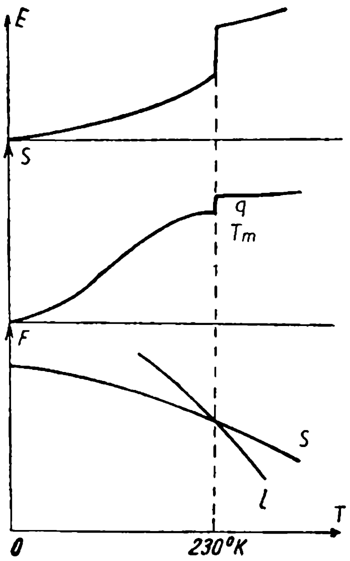


FIG. 190. Curves of internal energy, entropy and free energy for mercury.

The equilibrium state of a thermodynamic system is determined by the minimum of its free energy. The Gibbs free energy has the form

$$G = E + pv - TS. \qquad \ldots [11.2]$$

Changes in the free energy as a result of isothermal and isobaric processes $T =$ const. and $p =$ const. are expressed by the formula $dG = dE + p\delta v - TdS$. For solids and liquids the change of volume in isobaric processes is usually small (e.g. when a crystal melts the volume changes by 2-3%). Thus for condensed systems we can make use of the expression for the free energy in the Helmholtz form

$$F = E - TS. \qquad \ldots [11.2a]$$

By measuring the specific heat and the thermal content of a system, one can determine the temperature dependence of the entropy and the internal energy and free energy. In Fig. 190 are given curves of E, S, and F for a gram-molecule of mercury. These curves were obtained from calorimetric measurements. The jump in internal energy and entropy at 230° K, the temperature (T_m) at which mercury melts corresponds to the emission of the latent heat of melting. The free energy curve falls with increasing temperature on account of the increase of the entropy term. At $T = T_m$ this curve experiences a break as it consists of two branches, belonging respectively to the solid and liquid states and intersecting at the point with abscissa T_m.

Thermodynamics determines the state of the system using the mean thermodynamic parameters. Statistical physics enables us to treat the system statistically by considering the position and velocity of the particles at various moments of time. A given thermodynamic state must comprise an enormous number of dynamical states or *complexes*. The number of complexes determines the *probability* w of a given state.

From all possible states in a statistical system under given conditions, that state is realised which is the most probable and to which the greatest number of complexes belong. In an isolated system changes are possible from less probable to more probable states. States with maximum probability correspond to equilibrium of the system.

Entropy is connected with probability by the Boltzman formula $S = k \ln w$ from which it follows that an increase in entropy depends on an increase in probability.

2. The Free Energy of a Disordered Solid Solution

The Entropy of Mixing (Disorder)

Consider a crystal consisting of atoms A and B in amounts N_A and N_B. The total number of atoms $N = N_A + N_B$. In the disordered solid solution the atoms A and B occupy any sites of the crystal structure in a random manner.

Compare two states of a system at the same temperature: a disordered solid solution and a mixture of crystals of the pure components. The entropy of the solid solution is greater than the entropy of the mixture of crystals since in the former case there is a greater number of ways of distributing the atoms. The corresponding entropy defect is called the entropy of mixing or the entropy of disorder.

The number of ways of placing N atoms on N lattice sites in a solid solution is $w_1 = N!$. The number of ways of placing atoms in each of the crystals A and B is correspondingly $N_A!$ and $N_B!$. The total number of ways of distribution in the system consisting of a mixture of both crystals is $w_2 = N_A!\,N_B!$. The entropy of mixing S_m is then

$$S_m = S_1 - S_2 = k \ln w_1 - k \ln w_2 = k \ln \frac{N!}{N_A!\,N_B!}$$

Using Stirling's formula $\ln N_c = N \ln N - N$ which is very accurate, even when N is only ~ 10

$$\begin{aligned} S_m &= -k\,\{N_A \ln N_A + N_B \ln N_B - (N_A + N_B) \ln (N_A + N_B)\} = \\ &= -k\,[N_A \ln c + N_B \ln (1-c)], \end{aligned}$$

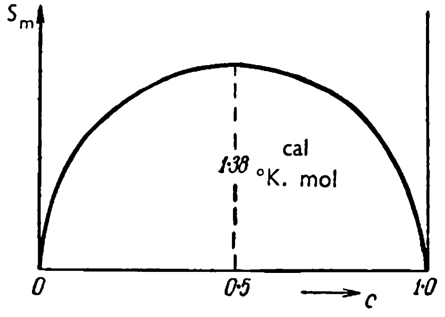


Fig. 191. Dependence of entropy on the concentration of a solid solution.

where $c = \frac{N_A}{N}$ is the concentration of component A. Hence

$$S_m = -kN[c \ln c + (1-c) \ln (1-c)]. \qquad \ldots [11.3]$$

Fig. 191 is a graph of the curve of S_m. This curve is symmetrical about the ordinate $c = 0{\cdot}5$. At $c = 0$ and $c = 1$ the curve is rising steeply. Such a shape of the curve explains the difficulty of preparing pure substances and removing small concentrations of impurities. When small amounts of impurity are introduced, free energy is reduced on account of the entropy term in $dF = dE - TdS$. The maximum value of the entropy of mixing is at $c = 0{\cdot}5$ which contributes to one mole $S_{max} = R \ln 2 = 1{\cdot}38$ cal/mol °K. a value comparable with those found for the entropy of solids.

The Internal Energy of a Solid Solution and the Energy of Mixing

The internal energy of a body at temperature T is $E = U + \int_0^T c_p dT$, where c_p is the specific heat and U the internal energy at $T = 0°$ K, that is the potential energy of formation of the solid solution from its atoms. To simplify the calculations we take account only of the interaction of nearest neighbours and assume that the energies of interaction of corresponding pairs of atoms u_{AA}, u_{BB}, u_{AB} are constants and do not depend on the type and position of neighbouring atoms. In this case the potential energy of a binary solid solution is

$$U = n_{AA}u_{AA} + n_{BB}u_{BB} + n_{AB}u_{AB}, \qquad \ldots [11.4]$$

where n_{AA}, n_{BB} and n_{AB} are the numbers of the various pairs of atoms.

In the disordered distribution of atoms the probability of finding atoms A or B on lattice sites is proportional to their concentrations. An arbitrarily chosen atom has zc neighbours of type A and $z(1-c)$ neighbours of type B, where z is the coordination number. Since $N_A = Nc$ and $N_B = N(1-c)$ the numbers of various pairs of atoms in the crystal are determined by the formulae

$$n_{AA} = \tfrac{1}{2} Nzc^2; \quad n_{BB} = \tfrac{1}{2} Nz(1-c)^2; \quad n_{AB} = Nzc(1-c). \qquad \ldots [11.5]$$

In calculating the pairs of the same kind n_{AA} and n_{BB}, each pair is counted twice; thus in the first two formulae we have to put in the coefficient $\frac{1}{2}$.

So, for the internal energy of a disordered solid solution we have, substituting equation [*5*] in [*4*]

$$U = \tfrac{1}{2} Nzc^2u_{AA} + \tfrac{1}{2} Nz(1-c)^2u_{BB} + Nzc(1-c)\,u_{AB} =$$
$$= \tfrac{1}{2} Nz\,[cu_{AA} - c(1-c)\,u_{AA} + (1-c)\,u_{BB} - c(1-c)\,u_{BB} + 2c(1-c)\,u_{AB}],$$

whence we obtain

$$U = \tfrac{1}{2} Nz\,[cu_{AA} + (1-c)\,u_{BB} + 2c(1-c)\,U_m], \qquad \ldots [11.6]$$

where U_m is the energy of mixing, given by

$$U_m = u_{AB} - \tfrac{1}{2}(u_{AA} + u_{BB}). \qquad \ldots [11.7]$$

The first two terms in formula [*6*] express the energy of an alloy consisting of a mixture of crystals of the pure components (for the case where the coordination numbers of both components are the same). These terms give a linear dependence of energy on concentration. The next term gives a parabolic dependence on concentration. Whether the energy of a solid solution is greater or less than the energy of a mixture of crystals of the components depends on the sign of the following term, which determines the sign of the energy of mixing. If the interaction of atoms of different kinds is large and exceeds the interaction of atoms of the same kind, then $u_{AB} < \frac{1}{2}(u_{AA}+u_{BB})$ and $U_m < 0$. This case favours the appearance of a solid solution or a chemical compound. If the interaction of atoms of different kinds is less than the interaction of atoms of the same kind then

$$u_{AB} > \tfrac{1}{2}(u_{AA}+u_{BB}) \text{ and } U_m > 0.$$

This case favours formation of a heterogeneous mixture of crystals.

The Free Energy of a Solid Solution

The entropy of a disordered solid solution is

$$S = \int_0^T \frac{c_p dT}{T} - R[c \ln c + (1-c)\ln(1-c)]. \qquad \ldots [11.8]$$

Here the first term is an expression for the vibrational part of the entropy due to random thermal motion; it is connected with heating of the body. The second term is the part of the entropy which is due to the configurational chaos and is the entropy of mixing.

For components with similar physical-chemical properties (this condition is necessary so that there may be a wide range of solubility) the specific heat c_p of the alloy does not depend much on composition. Thus in the expression for the free energy of a solid some of the terms depend only on temperature while other terms depend only on composition. Thus

$$F(c, T) = K(T) + U(c) + RT[c \ln c + (1-c)\ln(1-c)], \qquad \ldots [11.9]$$

where

$$K(T) = \int_0^T c_p dT - T\int_0^T \frac{c_p dT}{T}.$$

Fig. 192 shows in graphical form the internal energy U, entropy TS and free

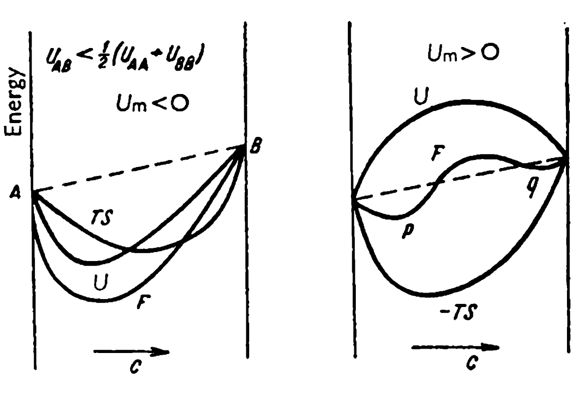


FIG. 192. Concentration dependence of the internal energy, entropy and free energy of a solid solution.

energy F of a solid solution as functions of composition for the following values of the energy of mixing: (*a*) $U_m < 0$ and (*b*) $U_m > 0$. Case (*a*) corresponds to the U-shaped curve of free energy with one minimum. The same shape for the curve of F also occurs for $U_m = 0$. Case (*b*) corresponds to a W-shaped curve for F with two minima p and q.

3. Equilibrium States of an Alloy

System with a U-shaped Free Energy Curve

Consider a system for which the free energy curve is U-shaped (Fig. 193). We require values for the free energy for various possible phase states of the alloy having composition c. If the system were a mixture of crystals of the pure components then the free energy of the alloys would change with concentration in an additive manner (straight line CD) and for concentration c we would obtain the point F. The formation of solid solutions leads to a reduction of free energy (points F_1 $F_2, \ldots$ obtained from the intersections of the lines C_1D_1, $C_2D_2, \ldots$). The smallest possible value of the free energy in this case (point F_3) corresponds to the intersection of the vertical ordinate CF with the free energy curve for the solid solution, and here the secant becomes a tangent to the curve F. It follows that for the U-shaped curve F, the equilibrium state of the alloy corresponds to a continuous series of solid solutions.

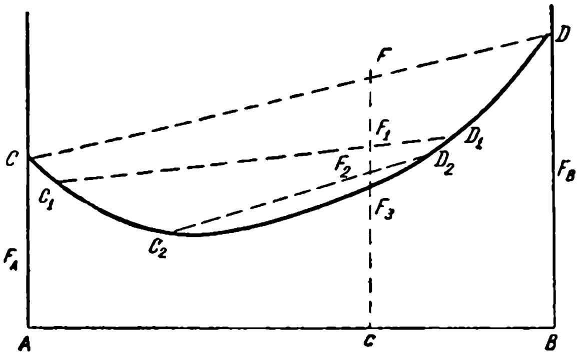


FIG. 193. Equilibrium state of an alloy with a U-shaped free energy curve.

System with a W-shaped Free Energy Curve

Consider now a system for which the free energy curve has the W-shape (Fig. 194). Take an alloy with concentration lying in the region of the intermediate maximum. In the case where the composition of the solid solution is retained, the value of the free energy is determined by the point F.

In a case where there is a breakdown into two terminal solid solutions α and β, the concentrations of which are determined by the points C_1 and D_1, the free energy of the heterogeneous mixture is determined by the point F_1, the

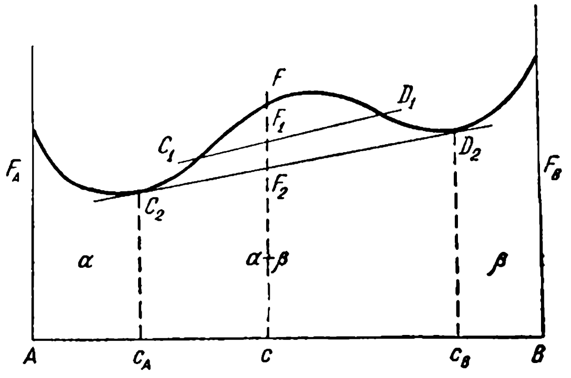


FIG. 194. Equilibrium state of an alloy with W-shaped free energy curve.

intersection of the straight line C_1D_1 with the vertical ordinate CF. Breakdown of the solid solution into two phases in this case leads to a lowering of free energy. Consequently the heterogeneous composition of the alloys is that of equilibrium whilst the homogeneous composition is unstable.

The minimum value of the free energy which results from the breakdown into two phases corresponds to the point F_2 which is the intersection of the vertical with the common tangent C_2D_2 to the two minima of the free energy curve. The limiting concentrations of solid solutions α and β, c_A and c_B are determined by perpendiculars dropped from the points of contact C_2 and D_2. In the range Ac_A the solution α is stable; in the range c_BB solid solution β is stable. The range c_Ac_B corresponds to a two-phase alloy consisting of a mixture of crystals of the solid solutions α and β with the limiting concentrations. The amount of each phase, α and β, is determined by the lever rule

$$\alpha = \frac{c_B - c}{c_B - c_A}; \quad \beta = \frac{c - c_A}{c_B - c_A}$$

System with Intermediate Phases

If in the ranges of concentration under consideration, there are not only solid solutions based on the pure components but also possible intermediate phases, then it follows that on the phase diagram we must consider low lying branches of the free energy curve for these phases (Fig. 195). By drawing the common tangents to the free energy curves we obtain the position of the one phase and two phase regions and in particular the regions of intermediate phases. The breadth of the ranges of solubility is determined by the shape and relative positions of the branches of the free energy curve. A sharp and narrow branch of the curve corresponds to a narrow region of solubility (phase γ).

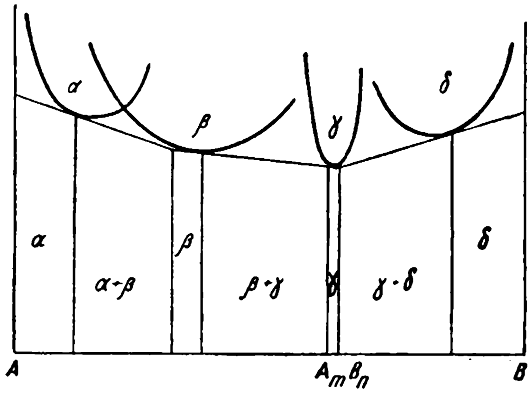


FIG. 195. Free energy curves for a system with intermediate phases.

Temperature Dependence of Phase Boundaries

In order to obtain the equilibrium diagram of some alloy we have to consider the free energy curves at various temperatures. As the temperature is lowered, the curves of free energy rise relative to the axis of abscissae, and the shape changes somewhat. Changes of shape of the curves and the rate at which they mix up is different for different phases (Fig. 196*a*). By drawing the common tangents at different temperatures $T_1 T_2 \dots$ we obtain the positions of the phase boundaries at those temperatures (Fig. 196*b*). By taking a sufficient number of temperature sections we can with accuracy construct a states diagram or equilibrium diagram.

4. Temperature Dependence of Solubility

Qualitative Considerations

In solid solutions, as in liquids, solubility usually increases with increasing temperature. In both cases this dependence has a common origin. Consider the equilibrium of a system which at low temperatures is a mixture of two phases. This will be the case if the internal energy of the mixture is less than that of the homogeneous solution (the case of Fig. 192*b*). At low temperatures the free energy $F = E - TS$ is approximately equal to the internal energy. As the temperature is raised the entropy term begins to make a substantial contribution to the free energy. This term increases with increasing temperature and has the effect of reducing the free energy. Since the entropy of a solid solution is greater than the entropy of a mixture of the pure components, a rise in temperature must be accompanied by a rise in solubility. In simple systems where there are no intermediate phases solubility always increases with increasing temperature. This is the case for example, in systems with unlimited solubility at high temperatures (Au – Ni, Au – Pt) and also in systems with limited solubility (Cu – Ag, Pb – Sn and others).

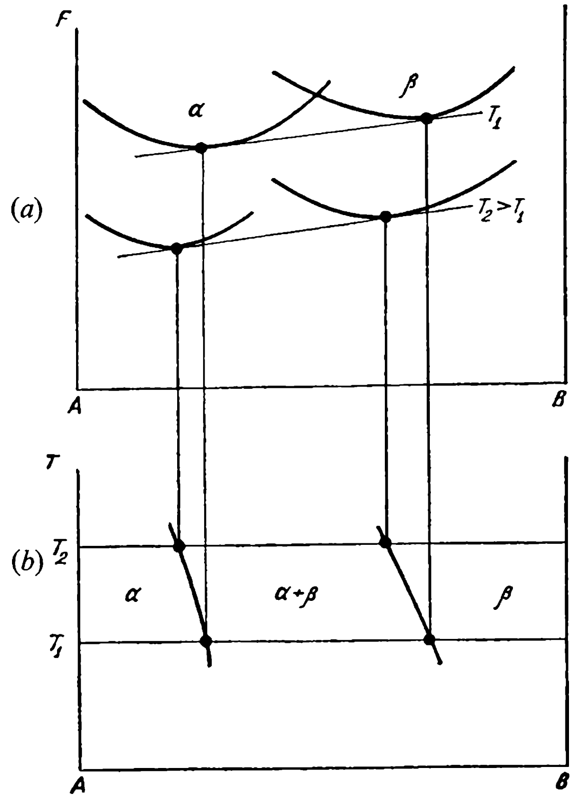


FIG. 196. Temperature dependence of solubility.

In systems where intermediate phases appear, these conclusions may be very different and the ranges of solubility of the individual phases may contract with increasing temperature. A well known example of such systems is found in the solubility curve for zinc in copper (α-brass). A qualitative explanation of such a solubility change is given in Fig. 196. This figure shows free energy curves for the initial solid solution α and for the intermediate phase β. If the β-phase is a disordered solid solution then its entropy of mixing is greater than that of solid solution α which has a smaller concentration. Thus as the temperature rises the free energy curve for the β-phase falls faster than does that for the α-phase. The slope of the common tangent thus falls with increasing temperature and this causes the point of contact of the tangent to move to the left, that is, into the region of lower concentration. As a result the solubility curves for both phases have the form shown in Fig. 196*b*. The range of solubility of the α-phase falls as the temperature rises while the solubility range of the β-phase rises in agreement with the general rule.

Equation of the Solubility Curve

The equation for the solubility curve can be derived from the expression for free energy. To simplify the calculations suppose $u_{AA} = u_{BB}$ and consider only those terms which depend on concentration. We then obtain from formula [*9*]

$$f = \frac{F}{N} = \text{const} + zU_m\{c(1-c) + p[c \ln c + (1-c)\ln(1-c)]\}, \qquad \ldots [11.10]$$

where f is the free energy referred to a single atom and $p = \dfrac{kT}{zU_m}$ is a non-dimensional parameter. Fig. 197 shows curves of f for various values of p. At low temperatures ($p = 0$) the system reduces to a mixture of the two components. As the temperature rises ($p > 0$) the curve takes on the W-shaped form and the positions of the points of contact of the common tangents to both parts of the curve become progressively closer. This shows that the solubility of both components increases with increasing temperature (Fig. 198). At $p = 0{\cdot}5$ the two minima come together and there is only one point of tangency. This corresponds to the merging of both branches of the solubility curves into the general dome-shaped curve. Above this curve there is continuous solubility. Below the curve the solid solution breaks down or the liquid solution emulsifies into two phases.

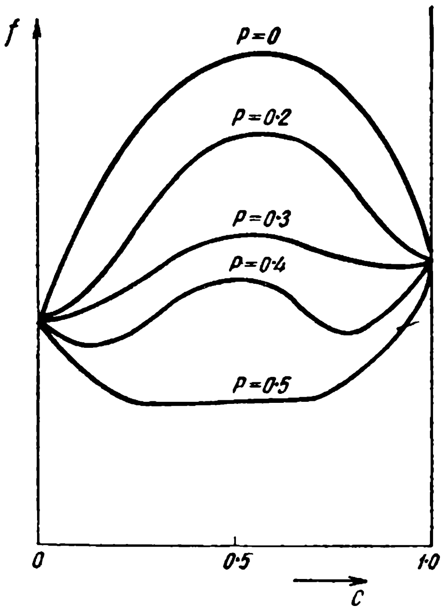


FIG. 197. Free energy curves for various values of p.

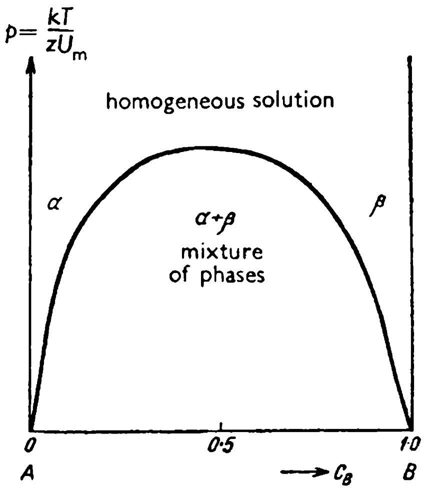


FIG. 198. Solubility curve for a binary system.

From the condition $\dfrac{df}{dc} = 0$ we obtain the equation of the solubility curve

$$1 - 2c + p(\ln c - \ln(1-c)] = 0,$$

and hence

$$\frac{c}{1-c} = e^{-\frac{zU_m(1-2c)}{kT}}. \qquad \ldots [11.11]$$

For small concentrations this gives an exponential form for the solubility curve

$$c \approx e^{-\frac{zU_m}{kT}},$$

and in many cases this agrees well with experiment.

5. Equilibrium Diagrams

We now give a thermodynamic interpretation of binary alloys based on their equilibrium diagrams. We assume for the liquid state that there is complete solubility of the components; that is that the free energy curve *l* is U-shaped for the liquid phase.

Continuous Series of Solid Solutions

In this case the atoms of both components have nearly the same radii, similar electronic configurations and the same crystal structure. The energy of mixing is nearly zero. The free energy curve for the solid phase *s* will also be U-shaped. At a high temperature T_1, curve *l* lies below curve *s* (Fig. 199*a*) since the entropy of the melt is greater than the entropy of the solid solution $S_l > S_s$. As the temperature is lowered curve *l* rises faster than curve *s* and at some temperature T_2 the curves come in contact. This moment corresponds to the onset of crystallization of the higher melting point component *A* (Fig. 199*b*). At a still lower temperature T_3 curves *s* and *l* cross (Fig. 199*c*). In accordance with the general rule we draw the common tangents to determine the concentrations c_s and c_l of the solid and liquid solutions which exist in equilibrium at that temperature. At temperature T_4, curve *l* has risen to such an extent that it makes its last contact with curve *s* (Fig. 199*d*). This corresponds to the end of crystallization of the lower melting point component *B*. At still lower temperatures curve *l* lies everywhere above curve *s*; this corresponds to equilibrium of the solid phase. By transferring the points of contact and the points of tangency on to a (*T*, *c*) diagram (Fig. 199*f*) we obtain the state diagram or equilibrium diagram of the alloy. This diagram is formed by the two lines of phase equilibrium, the liquidus line and the solidus line. These phase lines divide the whole diagram into three parts, the region of liquid solutions, the region of solid solutions and the heterogeneous region composed of a mixture of crystals and liquid. The last named region is lens- or cigar-shaped. Such equilibrium diagrams are found in the alloy systems Ag – Cu, Cu – Ni and Au – Pt.

However in the majority of systems where there is a continuous series of solid solutions the melting curves have an intermediate maximum or minimum (Fig. 200). A diagram with intermediate maximum arises in the case when the curve *s* has greater curvature than curve *l* and the first contact of the curves occurs in the intermediate range of concentration (Fig. 200*a*). If curve *s* has a smaller curvature than curve *l* then contact of the curves finishes in the intermediate region with the result that the diagram shows an intermediate minimum melting temperature (Fig. 200*b*).

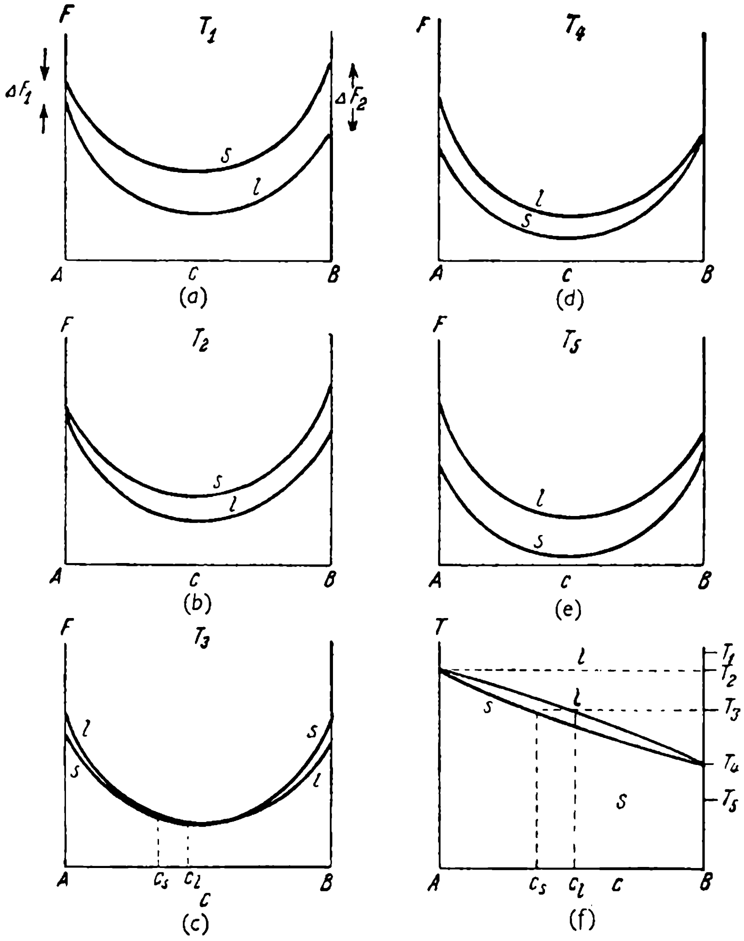


FIG. 199. Construction (*a*–*e*) of a lens-shaped equilibrium diagram (*f*) for a solid solution.

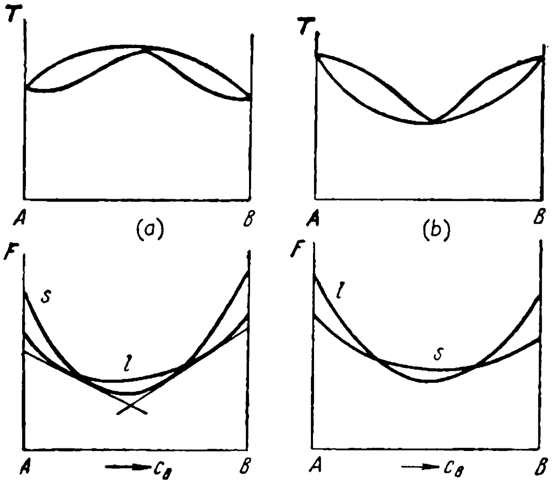


FIG. 200. Construction of an equilibrium diagram for solid solutions with maximum and minimum melting points.

The Eutectic

Suppose the curve s has two intermediate minima. This is possible in the first place for binary systems where the components can form a continuous series of solid solutions. In this case s is a continuous W-shaped curve (Fig. 201*a*). The second possibility is for binary systems where the components have different crystal structures and cannot form a continuous series of solid solutions. The free energy curve of such a system consists of two branches representing free energy curves of the two different solid solutions α and β based on components A and B (Fig. 202*a*). We can consider both these cases at the same time since they correspond to the same kind of equilibrium diagram.

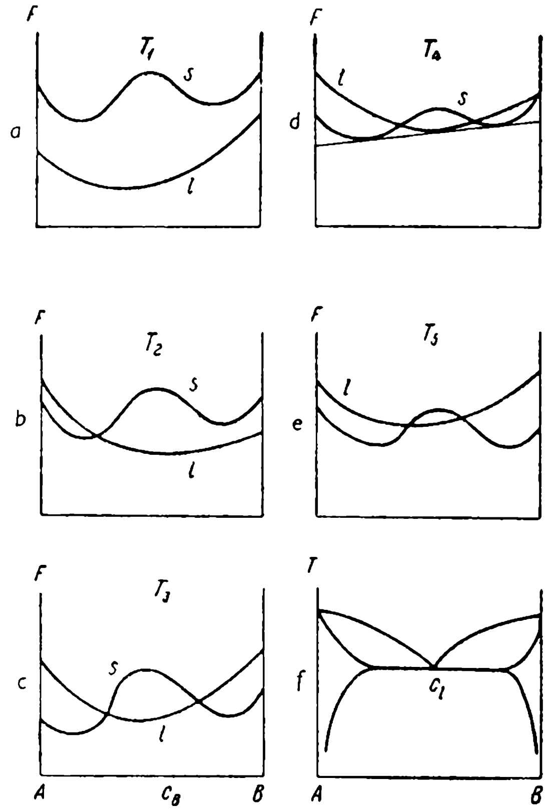


FIG. 201. Construction of an equilibrium diagram of the eutectic type.

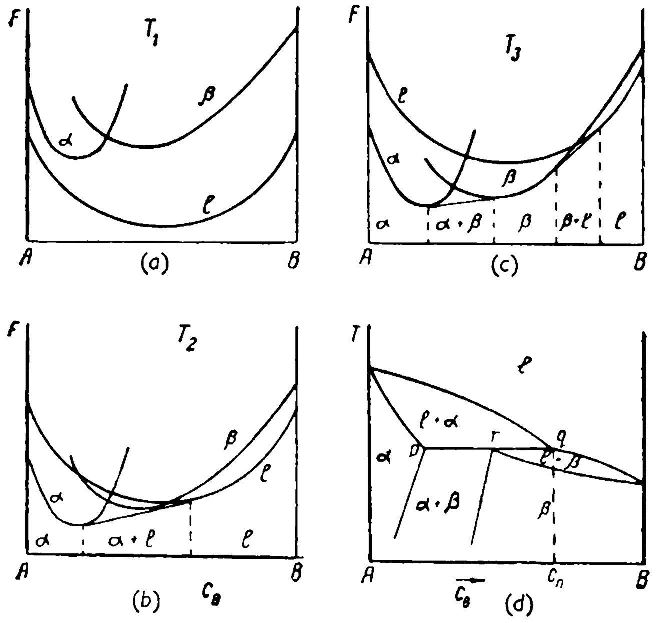


FIG. 202. Construction of an equilibrium diagram of the peretectic type.

Fig. 201*a-e* shows the relative positions of the curves *s* and *l* at various temperatures. At temperature T_4 the common tangent to the minima of curve *s* is also a tangent to curve *l* so that there is triple tangency. This temperature is the eutectic temperature and at it there are three phases in equilibrium; the liquid solution with eutectic concentration c_l and the two terminal solid solutions with limiting concentrations. The equilibrium diagram of the alloy (Fig. 201*f*) is the eutectic diagram. An example of a eutectic alloy where the components can form a continuous range of solid solutions, is the Cu – Ag system. Examples of eutectic alloys with dissimilar components are the alloys Cr – Ni, Al – Si, Pb – Sn, Pb – Sb etc.; this type is much more common.

The Peretectic

If the melting points of the components are very different, the equilibrium diagram of the alloy will have a somewhat different form. Fig. 202*a* shows the relative positions of curves of *s* and *l* at high temperatures. The difference between the free energies *s* and *l*, ΔF_A and ΔF_B is very different for the two components. Thus when we find threefold tangency, the point corresponding to equilibrium concentration of the liquid alloy c_n lies outside the limits of the region of solid phase equilibrium. This leads to a diagram of the peretectic type (Fig. 202*d*). This type of diagram is found with the alloys Ag – Pt and Cd – Hg.

Systems with an Intermediate Phase

Free energy curves at various temperatures for an alloy with one intermediate phase β are shown in Fig. 203*a-e*. Fig. 203*f* shows the corresponding equilibrium diagram. The intermediate phase β has a maximum in its melting temperature and crystallizes from the melt. The equilibrium diagram includes two eutectics. If one of the phases α or γ has a low melting point then the diagram will have a eutectic or a peretectic (Fig. 204). A second type of equilibrium diagram corresponds to the formation of an intermediate phase β by a peretectic reaction (Fig. 205*f*). The construction of such a diagram from the curve of free energy is shown in Fig. 205*a-e*. Formation of compounds by peretectic reactions takes place in many alloys—Bi – Ni, Na – K etc.

Complex Equilibrium Diagrams

Using the same thermodynamic principles as in the above simple types of equilibrium diagram we can understand the construction of more complex types. Complications in the equilibrium diagrams are connected with two circumstances: the occurrence of several intermediate phases in the system and phase transitions within the solid state. A large number of systems are known where there are several intermediate phases. In particular there are the alloys of the transition metals and the heavy monovalent metals with metals of sub-group B, where electron compounds are formed. The copper-zinc (brass) system is a typical example.

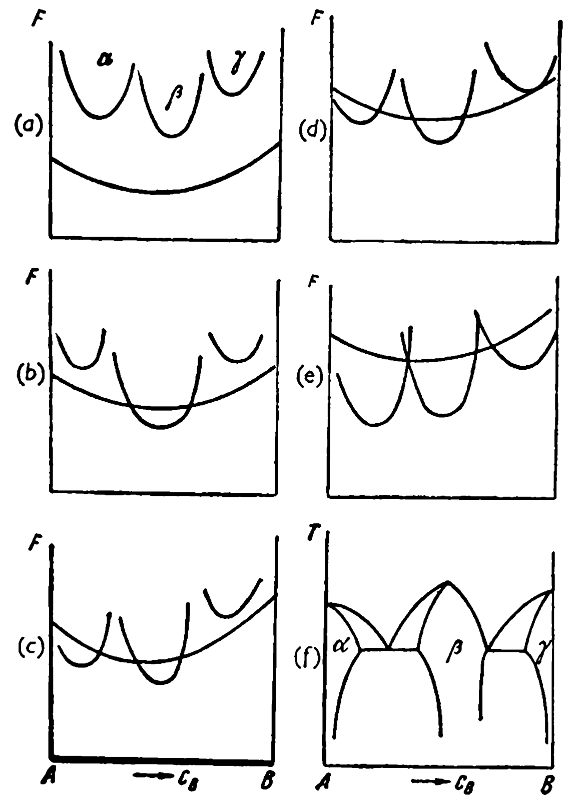


FIG. 203. Construction of an equilibrium diagram with an intermediate phase and two eutectics.

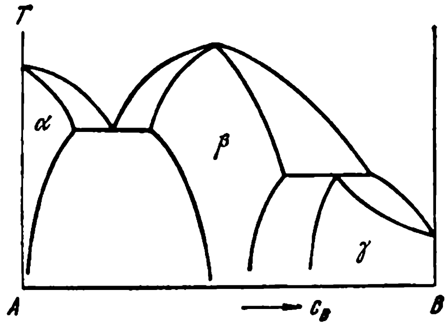


FIG. 204. An equilibrium diagram with an intermediate phase, an eutectic, and a peretectic.

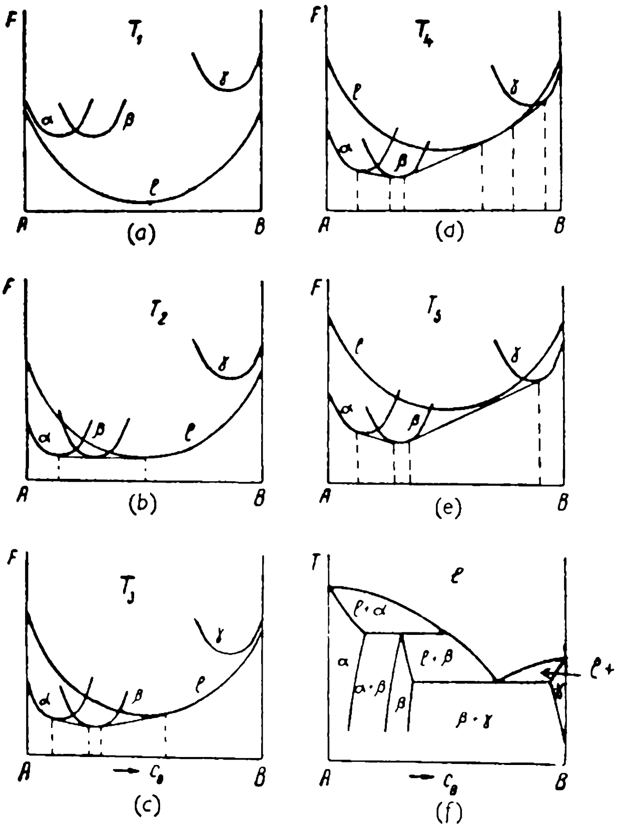


FIG. 205. Construction of an equilibrium diagram with peretectics.

6. Phase Transitions

Transitions within the Solid State

In thermodynamics, instability of phases leads to phase transitions. Instability of solid phases leads to the appearance of phase boundaries similar to those which occur in transitions from the liquid to the solid state for the thermodynamic laws of equilibrium which determine the character of phase curves do not depend on the state of aggregation. Curves of phase transitions called eutectoids and peretectoids are analogous to eutectics and peretectics which occur in the transition from the liquid to the solid state. A well known example of a eutectoid transition is the austenitic transition in the Fe–C system. Reduction of solubility in the solid state as the temperature falls, leads to breakdown of the solid solution and the formation of crystals of a new phase in a manner analogous to the formation of crystals from a supersaturated liquid solution.

In the solid state there may also be transitions from disorder to order occurring in a series of substitutional solid solutions with rational compositions, e.g. there are such transitions in the alloys CuAu, Cu_3Au, in β-brass and so on. These transitions are characteristic only of the solid state but, since the physical and chemical properties of the different phases are different, their recognition is essential for the practical utilization of the alloys. If we consider a case when the phase boundaries intersect as the temperature is changed then the phase state of the alloy must change. There are two basic cases where there is a transition from a single phase to a two-phase region; these are (*a*) phase boundaries inclined to the axis of abscissae (Fig. 206*a*) and (*b*) horizontal phase boundaries (Fig. 206*b*) such as occur in a eutectic or peretectic.

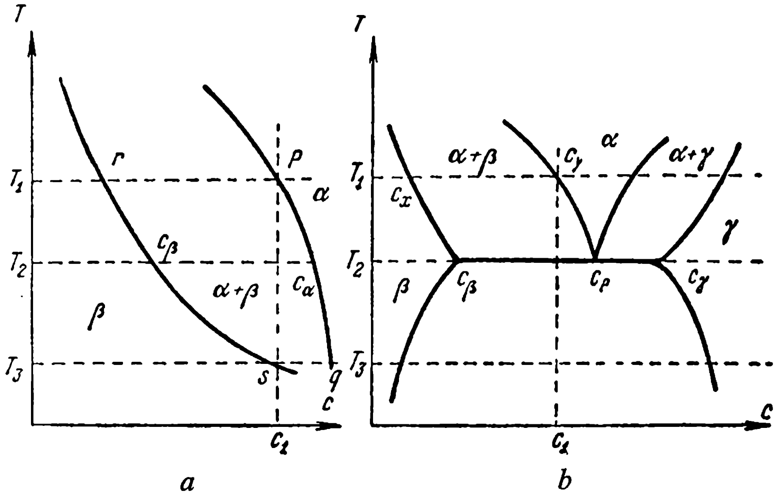


FIG. 206. Phase boundaries for case (*a*) breakdown of a solid solution and (*b*) eutectic breakdown.

Breakdown of a Supersaturated Solution

Consider an alloy of composition c_1 (Fig. 206*a*). At a temperature above T_1 the alloy is in the single-phase region α. Cooling below this temperature results in transition to the two-phase region $(\alpha+\beta)$. Finally at temperatures below T_3 the alloy comes into the single-phase region β. In the two-phase region at temperature T_2 the alloy consists of a mixture of crystals.

The amount of each phase is determined by the lever rule $\frac{\alpha}{\beta} = \frac{c-c_\beta}{c_\alpha - c}$ and is continuously changed as the cooling proceeds in such a way that the quantity of crystals of the α-phase falls while the amount of β-phase increases.

The composition of the phases also changes with temperature. The composition of the α-phase is determined by the line *pq* and that of the β-phase by the line *rs*. To begin with, crystals appear with composition markedly less than c_1 but towards the end of the phase transition, the composition is greater than c_1. The approach to equilibrium concentration requires redistribution of the atoms of the components and perhaps diffusion. In fact, equilibrium conditions are rarely attained and in consequence crystals are formed with mean concentration c_1 but not uniform in composition.

Such crystals have a zone structure; in the central parts which are formed first the concentration is lower than in the outer zones.

The Eutectic Transition

Fig. 206 includes a horizontal line of phase transition. If $T_1 > T_3$ this will represent a euctectic (or eutectoid) transition or if $T_3 > T_1$ it will represent a peretectic (or peretectoid) transition. At first during the transition to the two-phase region $(\alpha+\beta)$ primary crystals of β– appear from the α-phase, the phase transition proceeding as explained above. The amounts of the phases are determined by the ratio $\frac{\beta}{\alpha} = \frac{c_y - c}{c - c_x}$. At temperature T_2 the fraction of β crystals is $\frac{c_l - c}{c_l - \beta}$ and the fraction of α is $\frac{c-c_\beta}{c_l - c_\beta}$. Below the transition temperature T_2 the α-phase becomes unstable and breaks down into $\beta+\gamma$. In this way secondary crystals of β are formed. Close to the transition temperature the total amounts of the phases are given by the ratio $\frac{\beta}{\gamma} = \frac{c_\gamma - c}{c - c_\beta}$. Primary and secondary crystals can be distinguished by microscopical investigation. The secondary crystals of β which appear along with the crystals of γ at the temperature of the eutectic transition usually form a finely dispersed mixture with a structure characteristic of a eutectic. The total amount of eutectic in the alloy is equal to the amount of α-phase just before the moment of eutectic transformation.

$$\frac{\text{fraction of eutectic}}{\text{all alloy}} = \frac{c - c_\beta}{c_l - c_\beta}$$

The fraction of β crystals in the eutectic is $\frac{\beta}{\text{eutectic}} = \frac{c_\gamma - c_l}{c_\gamma - c_\beta}$ and does not depend on the concentration *c*.

Limitations of the Equilibrium Diagram

Equilibrium diagrams do not give all the characteristics of the state of a substance since by their nature they describe equilibrium compositions of a system and these may only be attained as a result of processes which take place very slowly. In practice, processes are completed with finite velocity and this leads to non-equilibrium conditions which do not agree with the equilibrium diagrams. We considered above the process whereby crystals are formed from a supersaturated solution, this leading to non-uniform composition within the boundaries of individual crystals of the new phase, in contradiction to the equilibrium diagram. Consider another example: a eutectic composed of crystals of $\beta+\gamma$. On heating above the eutectic temperature, crystals of α start to form as a result of interaction of crystals of γ with a part of the β-crystals, the mechanism being diffusional redistribution of the atoms of the components A and B. First crystals of α are formed on the surfaces dividing crystals of β and γ. This leads to spatial separation of the β and γ crystals and consequently to a slowing up of the transformation. With a finite rate of heating or cooling and insufficient time for diffusion processes this leads to the appearance of non-equilibrium states.

With high rates of heating which greatly exceed the rates of diffusion processes in solids it is possible to obtain metastable phases. These in general are not represented by equilibrium diagrams and do not occur on these diagrams. Examples of these are the martensitic phases occurring as a result of quenching in steel and various alloys. These metastable phases may survive for indefinitely long times and since they improve mechanical properties, are of great practical importance. In such cases one may construct diagrams describing metastable states as well as the usual diagrams of equilibrium states.

Another limitation of equilibrium diagrams is connected with the fact that they do not show the distribution of the crystals of the various phases—that is whether mixtures of coarse or fine grains, and the shapes, whether equi-axed, plate-shaped or needle-shaped. This limitation is specially significant since the shape and distribution of crystals is of great importance in that the mechanical properties are greatly influenced by these factors.

A fuller thermodynamic consideration of the equilibrium of alloys requires the addition of surface energy and the energy of internal stresses arising from phase transitions as a result of the difference in the volumes of crystals of the various sizes, and in their shapes and anisotropy. These factors are not reflected in the usual equilibrium diagrams. They lead to the existence of thermal hysteresis, that is to differences in the position of phase boundaries as observed on heating and cooling.

In conclusion, equilibrium diagrams do not reflect those changes of physical properties of a substance which arise from changes of thermodynamic parameters (temperature and pressure) within the limits of single phase regions. Such changes can be of great significance and because of their practical importance necessitate different technological treatments (for example, hot and cold deformation of metals). These changes are described by the general interdependence of composition, structure, and properties of materials which in the main is determined by the microscopical theory of matter.

Problems

1. Calculate the binding energy u_{AA} from the heat of sublimation L of a metal (copper).
2. Calculate the energy of mixing U_m of the gold-manganese alloy at the temperature of the maximum of the solubility curve $T_m = 1100°$ K.

Chapter 12

ORDERING AND DISORDERING OF THE ATOMIC STRUCTURE OF ALLOYS

1. The Ordered Structure of Substitutional Solid Solutions

Discovery of the Phenomenon of Ordering

In 1914 when Kurnakov and his co-workers were investigating the electrical resistivity of gold-copper alloys they encountered the difficulty of making them into wires. The cold work produced during drawing of wires with compositions close to Cu_3Au and CuAu was not reduced by annealing and the wires remained hard. This was the more unexpected because on the equilibrium diagram of the gold-copper system (Fig. 207) the liquidus and solidus curves indicated a continuous series of liquid and solid solutions. Detailed studies of the copper-gold system by physical and chemical methods showed that solid state phase transformations took place in the alloys at certain compositions. The curves of phase transformation outlined two dome-shaped regions. The temperature of the beginning of the transition (the Kurnakov temperature) was $T_k = 425°$ C for composition CuAu and $T_k = 395°$ C for Cu_3Au. These phase transitions were accompanied by large changes in the properties of the alloy and by thermal effects.

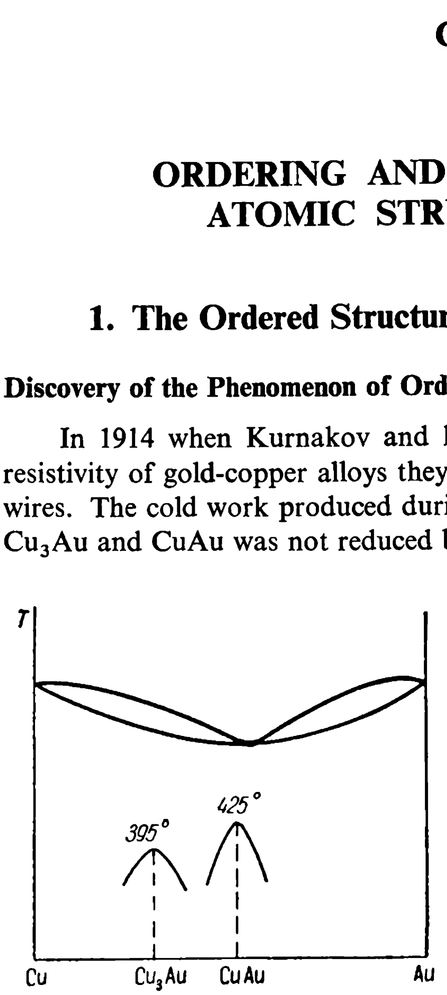


FIG. 207. Phase transformations in gold-copper alloys.

The Change in Electrical Resistivity

Fig. 157 shows the composition dependence of the electrical resistivity of quenched and annealed copper-gold alloys. The curve of $\rho(c)$ for quenched alloys has one smooth maximum; this is typical of a continuous series of solid solutions. The electrical resistivity of the solid solutions rises to 7-8 times that of the pure metals. In the case of the annealed alloys there are sharp minima corresponding to the rational compositions Cu_3Au and CuAu, these minima falling to values of electrical resistivity close to those for the pure metals.

Fig. 208 shows the temperature dependence of the electrical resistivity of the alloy Cu_3Au: a quenched alloy has the normal linear dependence on temperature.

An alloy which has been held for a long time at the temperature of measurement ('annealed') has a much lower electrical resistivity. A particularly sharp change takes place when the transition temperature T_k is reached. However the process of transition is not quite completed at that temperature but rather dies out gradually, stretching out over a large temperature interval of about 150-180°. At still lower temperatures the change of electrical resistivity of the annealed alloy once again follows a straight line parallel to that for the quenched alloy. One can draw some conclusions about the equilibrium of the intermediate states of the alloys from the fact that the curves obtained during heating and cooling almost coincide, showing only a small thermal hysteresis.

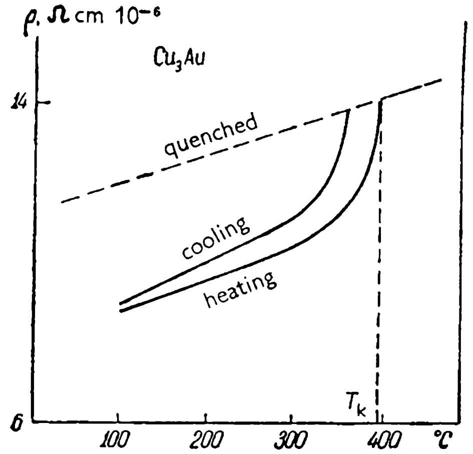


FIG. 208. Temperature dependence of the electrical resistivity of the alloy Cu_3Au.

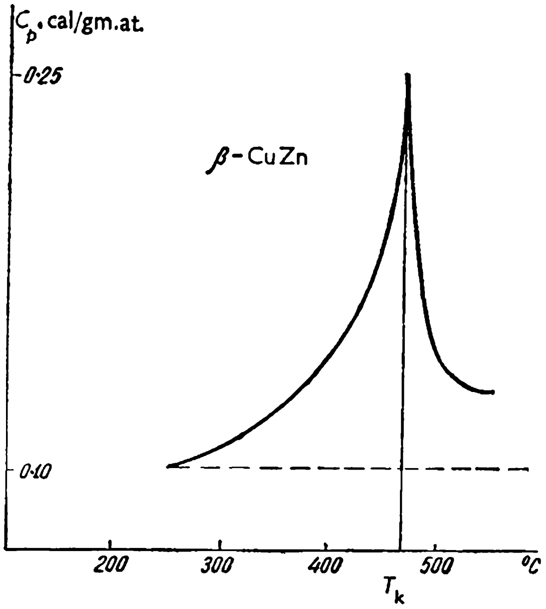


FIG. 209. Change in the specific heat of the alloy CuZn at the transition temperature.

Change of Specific Heat

Calorimetric measurements show that heat is given out during cooling through a phase transition and correspondingly absorbed during heating. The changes of heat content go on, not only at the temperature T_k but also throughout the whole temperature interval in which the transition is completed; as a result of this the specific heat of the alloy is increased in the region of the phase transition. Fig. 209 shows the variation in specific heat of the alloy β-CuZn. The curve has the characteristic λ-shape. At temperature T_k there is a discontinuity in slope of the specific heat. The absence of a latent heat enables us to attribute the β-brass transition to a second order phase transformation.

For Cu_3Au, the specific heat curve has a similar appearance but at the transition temperature the specific heat curve itself shows a discontinuity. This indicates that in this case there is latent heat and consequently a first order phase transition. In Fig. 209 the dotted line corresponds to the normal trend of specific heat which would occur in the absence of a transition, and judging by the slope of the c_p curve relative to this line the basic changes in the alloy take place below the transition temperature. However even above the transition the specific heat remains somewhat higher, showing that the transition may be incomplete and that there are still changes above the temperature T_k.

Structural Changes

The nature of the transition which takes place in the alloys we are considering was established by means of X-ray analysis. In the majority of cases the X-ray patterns of quenched and annealed alloys are very similar, indicating that the basic structural patterns are the same. The principal difference is that the X-ray photographs show weak additional lines indicating an increase of some lattice periodicities and of the volume of the unit cell by some small factor. The appearance of additional 'superlattice' lines shows that in the annealed specimens, an ordered arrangement of atoms on the lattice sites of the solid solution has been established in place of the disordered arrangement characteristic of the quenched alloys. The types of binary alloy structure were given in Chapter 6 §2.

Long Range Order

Between the limiting cases of completely ordered and completely disordered arrangement of the atoms A and B on the lattice sites α and β a continuum of transitional cases is possible corresponding to the structure with varying degrees of order. The appearance of superlattice lines shows that an ordered arrangement of atoms on the different α and β sites can be established without violating the periodicity of the arrangement, in volumes right up to the size of individual crystals. Such an arrangement of atoms in a crystal of a solid solution indicates the establishment of long-range order. A partially ordered arrangement can be quantitatively described by the *degree of long-range order* s which is linearly related to the probability of finding atoms on the correct positions and varies beteeen the limits $s = 0$ for the disordered structure and $s = 1$ for the completely ordered structure (§2).

Thermal and Concentration Disordering

The ordered arrangement of atoms corresponds to a lower internal energy of the alloy and is more stable at low temperatures. However the establishment of order requires thermal motion and diffusion of atoms within the crystal and the disordering effect of thermal motion increases with increase of temperature. Thus the degree of order depends on the temperature of the alloy. This degree of order s can be determined experimentally from the intensity of the superlattice lines. Fig. 210 shows changes with temperature in the degree of long range order: the process of thermal disordering gradually increases with temperature and is sharply accelerated as the temperature approaches T_k. For the alloy with composition AB there is a continuous transition while for composition A_3B there is a discontinuous change in the degree of order at the Kurnakov temperature.

When the composition of the solid solution departs from stoichiometry, the formation of a fully ordered periodic structure is not possible. Even with the greatest urge to ordering some of the excess atoms of one component must inevitably occupy sites belonging to the other. This leads to concentration disordering, that is to a lowering of the degree of order of the alloy structure as compared with the maximum possible in stoichiometric composition. Fig. 211 shows the changes in the degree of ordering in annealed copper-gold alloys as functions of concentration: the curves clearly show the large changes brought about by departures from rational composition. The smaller values of the ordering for the experimental curves (full lines) can be explained by the presence of some thermal disordering. The phenomenon of concentration disordering can also be seen on the curves giving the electrical resistivity of annealed alloys. (Fig. 157).

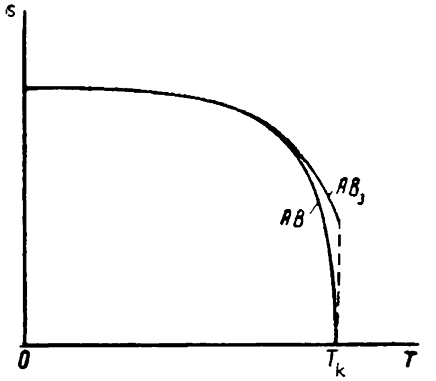


Fig. 210. Thermal disordering of alloys with compositions *AB* and A_3B.

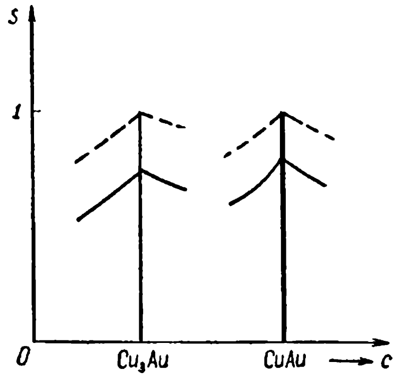


Fig. 211. Concentration disordering of alloys with compositions *AB* and A_3B. Dashed lines: theoretical; Full lines: experimental.

Short Range Order

The origin of ordering in solid solutions may lie in there being higher binding forces between atoms of different kinds than between atoms of the same kind. Thus atoms of one kind would seek to surround themselves with atoms of the other kind. This drive takes place at any temperature without regard to the temperature which determines whether there is long range order in a crystal or not. Moreover this drive takes place not only in the crystalline state but also in the liquid where the appearance of long range order is impossible in principle. Consequently the basis of ordering of atomic structure is the drive towards the establishment of short-range order. The degree of short range order (σ) can be determined from the number of different kinds of binding (see §3) in such a way that σ is equal to unity for a completely ordered state and to zero for a completely disordered state. The latter state is attained only in the limit of $T \to \infty$. At finite temperatures $\sigma > 0$. The establishment of long-range order is a secondary process such that the maximum value of σ is reached only when long-range order is established. The presence of short-range order enables us to explain the incompleteness of the process of ordering at the Kurnakov temperature and the residual disorder at temperatures above T_k.

2. Theory of Long-Range Order

The basis of the theory of long-range order was put forward by Gorsky and improved by Bragg and Williams and other workers. A more complete theory of the ordering of the structure of solid solutions based on the assumption of short-range order and including long-range order was put forward by Bethe and developed by Lifshits and others. The theory of order is beset with considerable difficulties, principally of a mathematical nature.

The Degree of Long-range Order and the Probability of an Atomic Distribution

We consider a substitutional solid solution (A, B) comprising two kinds of atoms A and B and distinguish two families of sites α and β in the lattice (α, β).

$$c_A+c_B = 1, \quad c_\alpha+c_\beta = 1, \qquad \ldots [12.1]$$

The total number of atoms N in a substitutional solid solution is equal to the total number of sites N'; $N = N'$. The numbers of atoms A and B are

$$N_A = c_A N; \quad N_B = c_B N. \qquad \ldots [12.2]$$

The numbers of sites α and β are:

$$N_\alpha = c_\alpha N; \quad N_\beta = c_\beta N. \qquad \ldots [12.3]$$

The concentrations of sites, c_α and c_β are determined by the numbers of sites α and β in the unit cell. Since the total number of sites in the unit cell of a metallic alloy is usually not large, the concentrations c_α and c_β are simple rational fractions. The two commonest cases are $c_\alpha = \frac{1}{2}$ and $c_\alpha = \frac{1}{4}$. The development of a completely ordered structure is possible only in these alloys where the number of atoms is equal to the number of sites.

$$N_A = N_\alpha, \quad N_B = N_\beta,$$

which leads to the relations

$$c_A = c_\alpha, \quad c_B = c_\beta, \qquad \ldots [12.4]$$

which determine the possible stoichiometric compositions of ordering alloys. In what follows we shall, without exception consider alloys with compositions which satisfy condition [*4*].

In the case of binary alloys it is sufficient to consider the distribution of atoms of one kind, e.g. A on the sites. If there is only one family of sites α and one family of sites β, the distribution of A atoms in the structure is uniquely determined by the probability of placing the atoms on sites of one family for example α; $p_\alpha^A = p$. The probability p for the limiting cases of order varies within the limits

$$\begin{aligned} p_{\min} &= c_A = c_\alpha \quad &&\text{(complete disorder)} \\ p_{\max} &= 1 &&\text{(complete order).} \end{aligned} \qquad \ldots [12.5]$$

As well as p we can use another parameter for the quantitative characteristic of ordering. This is the degree of long-range order s which is linearly related to the probability of the atomic distribution

$$s = \frac{p-p_{\min}}{p_{\max}-p_{\min}} = \frac{p-c_A}{1-c_A} = \frac{p-c_\alpha}{1-c_\alpha}. \qquad \ldots [12.6]$$

The degree of long-range order for alloys of various stoichiometric compositions varies within the limits from $s = 0$ to $s = 1$. By means of [*6*] we can express p as a function of s

$$p = p_{\min}+s\,(p_{\max}-p_{\min}) = c_\alpha+s\,(1-c_\alpha).$$

To obtain the distribution of atoms A and B on the sites α and β we note that N_α^A and N_β^A are the numbers of A atoms occupying sites α and β respectively, N_α^B and N_β^B

are the numbers of B atoms occupying sites α and β respectively. From conservation of numbers of sites and numbers of atoms it follows that

$$N_\alpha^A = N_\alpha p = c_\alpha \cdot p \cdot N; \; N_\beta^A = N_A - N_\alpha^A = (c_A - c_\alpha p)N = c_\alpha(1-p)N;$$
$$N_\beta^B = N_\beta - N_\beta^A = (1-c_\alpha)N - c_\alpha(1-p)N = (1-2c_\alpha + c_\alpha p)N;$$
$$N_\alpha^B = N_\alpha = N_\alpha^A - N_\beta^A. \qquad \ldots [12.7]$$

The last relation follows from the fact that any transfer of A atoms from α-sites to β-sites requires a transfer of B atoms from β to α.

The Number of Bonds and the Internal Energy of a Binary Solution

As shown previously (Ch. 11 §2) the principal contribution to the energy comes from the interaction of nearest neighbours. In a crystal of binary composition (A, B) there are bonds of three types AB, BB and AA. The total number of all bonds is

$$n = \tfrac{1}{2} Nz = n_{AA} + n_{BB} + n_{AB}, \qquad \ldots [12.8]$$

where n_{AA}, n_{BB}, n_{AB}, are the numbers of bonds of each type, z is the coordination number and N is the total number of atoms in the crystal. The half comes in since each bond is counted twice. We assume that the energy of the individual bonds u_{AA}, u_{BB}, and u_{AB} does not depend on the character of the surroundings and does not vary if the atoms are interchanged. The internal energies of crystals of the pure components are equal to their heats of sublimation at 0° K. Thus

$$U_A = nu_{AA} = \tfrac{1}{2} N_A z u_{AA}; \quad U_B = nu_{BB} = \tfrac{1}{2} N_B z u_{BB}. \qquad \ldots [12.9]$$

The internal energy of the solid solution is

$$U_{(A,\,B)} = n_{AA} u_{AA} + n_{BB} u_{BB} + n_{AB} u_{AB}. \qquad \ldots [12.10]$$

To evaluate this, we need to know the number of the corresponding bonds. Divide the crystal (A, B) by planes which pass through the bonds. To each of the ends of each bond cut in this way there belongs a $1/z$ part of the corresponding atom. The conditions for conservation of numbers of atoms A and B then give

$$c_A N = \left(\frac{2}{z}\right) n_{AA} + \left(\frac{1}{z}\right) n_{AB}; \quad c_B N = \left(\frac{2}{z}\right) n_{BB} + \left(\frac{1}{z}\right) n_{AB}.$$

Hence we obtain the general relations between the numbers of the various types of bond in a two component crystal

$$n_{AA} = c_A n - \tfrac{1}{2} n_{AB}; \quad n_{BB} = c_B n - \tfrac{1}{2} n_{AB}. \qquad \ldots [12.11]$$

Putting [*11*] into [*10*] we obtain

$$U_{(A,\,B)} = c_A n u_{AA} + c_B n u_{BB} + n_{AB}\,[u_{AB} - \tfrac{1}{2}(u_{AA} + u_{BB})],$$

which can be written in the form

$$U_{(A,\,B)} = U_A + U_B + n_{AB} \,.\, U_1, \qquad \ldots [12.12]$$

where

$$U_1 = u_{AB} - \tfrac{1}{2}(u_{AA} + u_{BB}). \qquad \ldots [12.13]$$

The first two terms in [*12*] are constant and do not depend on the distribution of atoms A and B on the lattice sites of the solid solution. They are numerically equal to the energies which crystals of the pure components would have if they were picked out from the solid solution. The third term $n_{AB}\, U_1$ is variable and depends on the atomic distribution. Formula [*12*] shows that the total energy of a solid solution $U_{(AB)}$ depends on the magnitude of the energy of all three types of bond u_{AA}, u_{BB} and u_{AB}, while the change of energy of a solid solution arising from interchange of atoms (configurational energy)

$$U_{(AB)}-(U_A+U_B) = n_{AB}U_1$$

depends only on one quantity U_1—the energy of mixing, which is equal to the change in energy of a solid solution arising from a change by one in the number of pairs of different kinds: $\Delta n_{AB} = 1$. The configurational energy of a solid solution thus changes by multiples of U_1 and reaches a minimum ($U_1 < 0$) when n_{AB} is a maximum corresponding to maximum ordering.

We now investigate the number of bonds n_{AB} in the case where there is a departure from stoichiometric composition on the assumption that the alloy retains its maximum possible order. Take an alloy of the CuZn type in which there is a chessboard alternation of α and β sites and consider the concentration range $0 \leqslant c_A \leqslant \frac{1}{2}$ (if $c_A > \frac{1}{2}$ we interchange atoms A and B). In the ordered state all β sites are occupied by B atoms and α sites are occupied by A and B atoms. Each A atom is surrounded by B atoms only. Thus

$$n_{AB} = N_A z = c_A N z = 2nc_A. \qquad \ldots [12.14]$$

The curve of n_{AB} consists of two straight line branches lying symmetrically with respect to the ordinate $c = \frac{1}{2}$ and giving a sharp maximum at their intersection. In Table 63 are set out the results of calculations for other types of solid solution. The curve of $n_{AB}(c)$ is symmetrical for an alloy of the CuAu type and unsymmetrical for an alloy of Cu_3Au type. The shape of $n_{AB}(c)$ curves in general corresponds to that of $s(c)$ curves (Fig. 211).

Table 63

Dependence of n_{AB} on c_A and s; $\nu_0^* = (\delta n_{AB})_0$

Structural type	$n_{AB}(c)$	$n_{AB}(s)$	ν_0
CuZn	$2nc_A$	$\frac{1}{2}n(1+s^2)$	16
CuAu	$\frac{2}{3}c_A(3-2c_A)$	$\frac{1}{6}n(3+s^2)$	8
Cu_3Au	$\frac{8}{3}nc_A(1-c_A)$ $(c_A > \frac{1}{4})$ $2nc_A\ (c_A < \frac{1}{4})$	$\frac{1}{8}n(3+s^2)$	4

* For a determination of ν_0 see formula (18).

Let us now investigate the dependence of the number of bonds n_{AB} on the degree of order in an alloy of the CuZn type. Consider the family of α-sites which are occupied by A and B atoms with respective probabilities p and $1-p$. In the case of α-sites occupied by A atoms we have to calculate the number of B atoms which occupy neighbouring β-sites. For the α-sites occupied by B atoms we have to calculate the number of neighbouring A atoms. Hence the number of bonds between atoms of different kinds is

$$n_{AB} = N_\alpha pzp + N_\alpha (1-p)\, z\, (1-p) = \tfrac{1}{2} Nz\, [p^2 + (1-p)^2] = \\ = \tfrac{1}{2} n\, (1+s^2). \qquad \ldots [12.15]$$

The results of analogous calculations for various types of solid solution are given in Table 63.

In all cases n_{AB} depends on the square of the degree of order.

The Energy of Ordering

For these cases considered in Table 63, a change in the configurational energy (energy of ordering) connected with a change in the degree of ordering, depends linearly on s

$$-\frac{dU_{(AB)}}{ds} = U_1 \frac{dn_{AB}}{ds} = \gamma U_1 ns. \qquad \ldots [12.16]$$

Here γ is a numerical quantity the magnitude of which depends on the structural type.

Denote by w the change of energy of a crystal produced by interchanging one pair of atoms A and B which were previously occupying their correct sites (w is thus an energy of disordering). Assuming that w is positive the energy of disordering must be supplied to the crystal and transfers it to a state of higher energy. If we exclude fluctuations in the distribution of A and B atoms and take a mean value of w for the whole crystal, then it follows from formula [*16*] that w must depend linearly on s

$$w = w_0 \cdot s. \qquad \ldots [12.17]$$

The energy of disordering has its maximum value when pairs of A and B atoms are interchanged under conditions of complete order $w(1) = w_0$. For a completely disordered structure $w(0) = 0$ for in this case there is no difference between sites α and β and interchange of any pair of atoms does not change the energy state of the crystal. In accordance with the relation

$$w = -\delta n_{AB} U_1 s = -\nu U_1 s, \quad (U_1 < 0;\ w > 0),$$

where $\delta n_{AB} = \nu$ (the change in the number of bonds between atoms of different kinds produced by interchanging two non-adjacent atoms A and B) we have that the maximum value for w is

$$w = -\nu_0 U_1. \qquad \ldots [12.18]$$

Here ν_0 is the largest change in the number of bonds between atoms of different kinds which takes place when two atoms A and B are interchanged. We can find ν_0 if we know the coordination ratios in ordered and disordered structures (Table 34, pp. 182-3). Table 63 gives values of ν_0 for three kinds of solid solution with AB structures and chessboard-type alternation of α and β sites.

Thermodynamic Theory of Long-range Order

An ordered arrangement of atoms arises from a disordered arrangement because it is more favourable for A atoms to be on α-sites than for them to be on β-sites. Transfer of an A atom from an α-site to a β-site is accompanied by the transfer of a

B atom from a β-site to an α-site and requires the expenditure of energy w. In an alloy at temperature T an equilibrium distribution of atoms on their own and on foreign sites is set up. This distribution is characterised by a quantity s, which corresponds to the difference $w(s)$ in energy between sites α and β. Let us investigate the distribution of atoms of a solid solution on their own and on foreign sites by applying Boltzmann's principle. The population of foreign sites β (meaning the ratio of the numbers of occupied and unoccupied sites) is equal to the population of their own α-sites multiplied by the Boltzmann factor

$$\frac{N_\beta^A}{N_\beta^B} = \frac{N_\alpha^A}{N_\alpha^B} \cdot e^{-\frac{w}{kT}}, \qquad \ldots [12.19]$$

whence

$$\frac{N_\alpha^A N_\beta^B}{(N_\beta^A)^2} = e^{\frac{w}{kT}}.$$

Introducing formula [*7*] we get

$$\frac{p(1-2c_\alpha-c_\alpha p)}{c_\alpha(1-p)^2} = e^{\frac{w}{kT}}. \qquad \ldots [12.20]$$

Expressing p as a function of s and putting that in [*20*] gives

$$\frac{(c_\alpha+c_\beta s)(c_\beta+c_\alpha s)}{c_\alpha c_\beta(1-s)^2} = e^{\frac{w}{kT}}. \qquad \ldots [12.21]$$

Equation [*21*] gives the relationship between the degree of order and the temperature of an alloy.

That equation [*21*] determines the equilibrium state of an alloy follows from the condition of minimum free energy

$$\frac{dF}{ds} = \frac{dU}{ds} - T\frac{dS}{ds} = 0, \qquad \ldots [12.22]$$

where U is the configurational energy of the alloy and S is the configurational part of the entropy. The internal energy due to disordering is related to the interchange of dN_α^A atoms from α-sites to β-sites and is given by

$$dU = -w dN_\alpha^A = -w c_\alpha N dp = -w(s) c_\alpha N(1-c_\alpha) ds, \qquad \ldots [12.23]$$

where U is the energy of disordering required to interchange one pair of A and B atoms. The configurational part of the entropy of a binary solid solution can be determined by the formula

$$S = k \ln W = k \ln (W_\alpha \cdot W_\beta), \qquad \ldots [12.24]$$

where W_α and W_β are the numbers of independent permutations of A and B atoms on α and β sites.

$$W_\alpha = \frac{N_\alpha}{N_\alpha^A!\,N_\alpha^B!}; \qquad W_\beta = \frac{N_\beta!}{N_\beta^B!\,N_\beta^A!}. \qquad \ldots [12.25]$$

Putting [*25*] into [*24*] and using Stirling's formula we obtain after some rearrangement

$$S(p) = C - Nk\,[c_\alpha p \ln p + (1-2c_\alpha + c_\alpha p)\ln(1-2c_\alpha+c_\alpha p) + \\ + 2c_\alpha(1-p)\ln(1-p) - c_\alpha p \ln c_\alpha], \qquad \ldots [12.26]$$

where C denotes those terms which do not depend on p. Differentiation of [*26*] gives

$$\frac{dS}{ds} = \frac{dS}{dp}\cdot\frac{dp}{ds} = (1-c_\alpha)\frac{dS}{dp} = -(1-c_\alpha)c_\alpha Nk \ln\frac{p(1-2c_\alpha+c_\alpha p)}{c_\alpha(1-p)^2}. \qquad \ldots [12.27]$$

Substitution of [*23*] and [*27*] in the equation of equilibrium leads to the equation

$$\ln\frac{p(1-2c_\alpha+c_\alpha p)}{c_\alpha(1-p)^2} = \frac{w}{kT},$$

which agrees with equation [*20*] which we obtained before by means of the Boltzmann relation.

Consider the graphical solution of equation [*21*]; take logarithms of the left and right hand sides:

$$L(s) = \ln\frac{(c_\alpha + c_\beta s)(c_\beta + c_\alpha s)}{c_\alpha c_\beta (1-s)^2}, \qquad \ldots [12.28]$$

$$P(s) = \frac{w}{kT} = \frac{w_0 s}{kT}. \qquad \ldots [12.29]$$

$L(s)$ is a function which increases with increasing s since the argument of the logarithm [*28*] is an increasing function of s. $L(0) = 0$ and $L(1) = \infty$. Thus the graph of $L(s)$ is a curve through the origin and tending asymptotically to indefinitely large values as s approximates to $s = 1$ (Fig. 212). $P(s)$ is a straight line through the origin with slope depending on temperature. As the temperature increases the gradient of the line is reduced and in the limit of $T \to \infty$ the line becomes coincident with the axis of abscissae. At high temperatures equation [*2*] has only one solution $s = 0$ which corresponds to the disordered state. At the absolute zero of temperature the straight line is coincident with the ordinate axis, while at very low temperatures there are two solutions $s_1 = 0$ and $s_2 = 1$ of which the second corresponds to a completely ordered structure. This solution determines the equilibrium state of an alloy since there corresponds to it a minimum of free energy which at $T = 0$ coincides with the minimum potential energy of the alloy.

Consider the solution of equation [*21*] for two cases (*a*) an alloy of symmetric composition AB ($c_\alpha = c_\beta = \frac{1}{2}$) and (*b*) an alloy of composition AB_n for example AB_3 ($c_\alpha = \frac{1}{4} \neq c_\beta = \frac{3}{4}$). For these cases the equation for $L(s)$ takes the forms

$$(a)\quad L(s) = 2\ln\frac{1+s}{1-s}, \qquad (b)\quad L(s) = \ln\frac{(1+3s)(3+s)}{3(1-s)^2}. \qquad \ldots [12.30]$$

Figs. 212 and 213 show curves of $L(s)$ and of the lines $P(s)$ which cut them. In the general case of $c_A \neq c_B$, the curve of $L(s)$ has a point of inflection. In the symmetrical case $c_A = c_B$, the curve of $L(s)$ has no point of inflection. In case (*b*), there are some temperatures for which the line P_3 intersects the curve L at three points, 0, Q and R and in case (*a*) at two points 0 and R. Analysis shows that in both cases the highest point R corresponds to a minimum of free energy so R determines the equilibrium value of the degree of order at temperature T. As the temperature rises the line P

falls as the point R approaches the origin, corresponding to a lowering of the degree of order. At some temperature $T = T_k$ the line becomes a tangent P_2 with point of tangency K for alloy AB_n and the origin for alloy AB. All points of curve L for the alloy AB represent possible equilibrium states, the degree of order falling from $s = 1$ at $T = 0$ to $s = 0$ at $T = T_k$. The rate of change of s increases as T approaches T_k. In the alloy AB_n the upper part of the curve L beginning at K corresponds to an unstable state. At temperature $T = T_k$ a limiting value of the degree of order $s = s_k$ is reached after which further lowering of the temperature leads to a discontinuous change of order down to the value $s = 0$. At temperature T_k the fully ordered alloy exists in equilibrium with an alloy which has the limiting degree of order s_k.

The solution of equation [*21*] results in a curve $s = s(T)$ which is, in the main, in agreement with the experimental curves (Fig. 210). A discontinuous change of order requires that a finite amount of energy be supplied to disorder the alloy. Thus the transition in the alloy AB_n is accompanied by a latent heat and is a first order phase transition. In the alloy AB with symmetrical composition the degree of order changes continuously; the corresponding transition takes place without a discontinuous change of energy and is a second order phase transition.

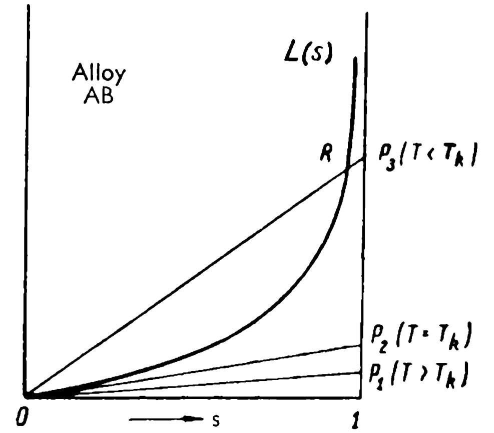


FIG. 212. Graphical solution of equation (*29*) for an alloy of composition AB.

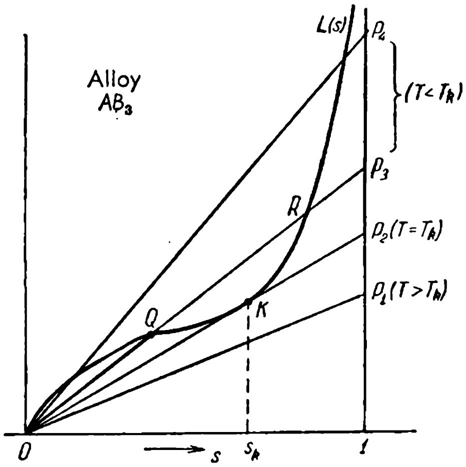


FIG. 213. Graphical solution of equation (*29*) for an alloy of composition AB_3.

For an alloy of the CuZn type $w = w_0 s = -16\, U_1 s$, so equation [*21*] has the form

$$\ln\frac{1+s}{1-s} = -\frac{8U_1}{kT}\, s. \qquad \ldots [12.31]$$

If ϕ_0 is the slope of the tangent to the curve $\ln\dfrac{1+s}{1-s}$ at the origin, then $\tan\phi_0 = 2$ and equating this to the slope of the limiting line P_2 (see Fig. 212) we obtain an expression for the Curie temperature or Kurnakov temperature,

$$T_k = -\frac{4U_1}{k}. \qquad \ldots [12.32]$$

Using this we can write equation [*31*] in the form

$$\ln\frac{1+s}{1-s} = \frac{2T_k}{T}\, s. \qquad \ldots [12.33]$$

At temperature $T<T_k$ the line $P(s)$ cuts the curve $L(s)$ at two points $s_1 = 0$ and $s_2 = s(T)$. The second derivative of the free energy

$$\frac{d^2F}{ds^2} = kN\left(\frac{T}{1-s^2}-T_k\right)$$

is negative at $s_1 = 0$ and positive at $s_2 = s(T)>0$. Thus the minimum of free energy corresponds to the second solution which gives the temperature dependence of the degree of order at equilibrium.

From formula [*26*] we see that the change of entropy produced by complete disordering is $\Delta S = R\ln 2 = 1{\cdot}38$ cal/mole. °K while for β-CuZn experiment gives $\Delta S = 1{\cdot}01$ cal/mole. °K a value close to the theoretical one but somewhat less than it. The total energy of disordering is

$$\Delta U = \int_0^1 dU = -4NU_1\int_0^1 sds = -2NU_1 = \tfrac{1}{2}RT_k.$$

For β-brass $T_k = 743°$ K, giving $\Delta U_{\text{theor}} = 740$ cal/mole while calorimetric measure. ments give $\Delta U_{\text{meas}} = 630$ cal/mole, again somewhat lower than the theoretical value- The energy expended in disordering is manifested by an increased value of the specific heat. The additional part of the specific heat which is due to disordering is equal to

$$\Delta C_v = \frac{dU}{dT} = \frac{dU}{ds}\cdot\frac{ds}{dT} = 4NU_1 s\frac{ds}{dT} = -Rs\frac{ds}{d\left(\frac{T}{T_k}\right)}.$$

The trend of specific heat of disordering with temperature can be obtained graphically by differentiating the curve $s = s(T)$.

In its essentials the theoretical curve agrees with experiment but there are the following differences: from theory it follows that the process of transition takes the whole temperature interval from 0° K to T_k while experiment shows that the transition takes place mainly at high temperatures. Thus the experimental curve has a greater tendency towards the λ-shape than the theoretical; this can be explained by the reduction of the diffusional mobility of atoms at low temperatures. The most significant difference between the curves is in the temperature range $T>T_k$. Here the experimental curve reveals an enhanced value of the specific heat which suggests that the process of disordering is still going on at temperatures above T_k. This is a further illustration of a difference between theory and experiment which was mentioned above in connection with the changes of entropy and energy of disordering. At temperatures above T_k, short-range order is preserved and the degree of this falls continuously as the temperature rises. The decay of short-range order is accompanied by a change of entropy of the alloy and requires the expenditure of energy.

3. Short-Range Order

The Degree of Short-range Order

Although the theory of long-range order explains some of the basic phenomen- of ordering, it does not explain the properties of an alloy above the transition team perature. The failings of the theory of long-range order are in matters of principle.

The drive for the formation of bonds between atoms of different kinds under the condition $U_1 < 0$, takes place in the alloy at any temperature. The number of these bonds depends on temperature and falls as the temperature is raised. However more or less large regions of order survive at any temperature. The basis of the phenomenon of ordering is really short-range order and the most complete and most recent theory of ordering is one based on the idea of short-range order since the forces of interatomic interaction fall off rapidly as the distance between atoms is increased. Thus in the favoured theories of ordering the interactions between nearest neighbours only are taken into account, this being a reasonable assumption and sufficient to give short-range order. However this assumption cannot explain how a state of long-range order arises starting from a disordered state. Moreover long-range order presupposes that there is a division of all the lattice sites, including those sites far from a given atom, into two families and each of those families is preferentially occupied by atoms of one kind throughout the whole volume of the crystal. In a less obvious form such a consideration involves the assumption that the energy of a given atom is essentially determined by its interaction with all the other atoms in the crystal. One may suppose that in a more complete theory of ordering, long-range order would arise on the basis of short-range order in the form of states having a very high degree of short-range order.

Let us define q the parameter measuring short-range order by the ratio of the numbers of bonds between atoms of different kinds to the total number of all bonds in the crystal

$$q = \frac{n_{AB}}{n} = \frac{n_{AB}}{n_{AA}+n_{BB}+n_{AB}}. \qquad \ldots [12.34]$$

The parameter q varies between the limits $q_{\min}$ for a completely disordered state to $q_{\max}$ for a completely ordered state. We represent the degree of short-range order by the quantity

$$\sigma = \frac{q-q_{\min}}{q_{\max}-q_{\min}}, \qquad \ldots [12.35]$$

which varies between the limits $\sigma \leqslant o \leqslant 1$.

The Ideas of the Theory of Short-range Order

Consider an alloy with composition AB with a chess-board distribution of sites. In this case

$$(n_{AB})_{\max} = n, \quad (n_{AB})_{\min} = \tfrac{1}{2} n,$$

and formula [35] becomes

$$\sigma = 2\left(q-\tfrac{1}{2}\right). \qquad \ldots [12.36]$$

Hence we can obtain the dependence of the number of unlike bonds on the degree of short-range order

$$n_{AB} = \tfrac{1}{2} n\,(1+\sigma). \qquad \ldots [12.37]$$

This expression is a linear function of σ. In a state which is almost completely ordered there must be a relation between the criteria for short- and long-range order since in the limiting case, $\sigma = s = 1$ (Fig. 214*a*) represents the same state of complete order of the structure. Equating formulae [*15*] and [*37*] gives the relation between σ and s:

$$\sigma = s^2 \quad (\sigma \geqslant 0,\ s \geqslant 0). \qquad \ldots [12.38]$$

An alloy may have a high degree of order while at the same time an estimate based on the degree of long-range order shows the absence of order. In the case shown in Fig. 214*b* the whole crystal consists of completely ordered regions which are mutually incoherent and are called antiphase regions. The boundaries between antiphase regions pass between pairs of like atoms and interrupt the strict periodicity of the crystal. If positive and negative antiphase regions in a crystal occupy equal volumes then in effect $s = 0$. At the same time the physical properties of such a crystal are nearly those of a crystal with a high degree of order. The criterion of short-range order better reflects the mean state of order and gives in this case $\sigma = 0{\cdot}7$. Fig. 214*c* shows some intermediate state of order for which relation [*38*] is sufficiently well satisfied.

The basic conclusions of the theory of short-range order for alloys of composition *AB* are as follows: the degree of short-range order σ decreases steadily as the temperature rises but never reaches zero. There is no critical temperature for σ similar to the Kurnakov temperature. This result is of the theory in agreement with experiment which shows that there is short-range order at temperatures above T_k. The new theory further sees the establishment of long-range order as a process analogous to crystallization in the liquid and the subsequent growth of the crystals which form. At temperatures above T_k there are in the alloy, antiphase regions of short-range order existing in dynamic equilibrium. Thermal motion and atomic diffusion continually break up these regions while the interaction between unlike atoms provides a mechanism for the establishment of order. Analogous processes take place in the melt above the temperature of crystallization, when ordered regions of atoms continuously arise and break up, their structure being nearly crystalline (atomic 'swarms'). At the crystallization temperature it becomes possible to form ordered groups of atoms which are the nuclei of the new crystalline phase.

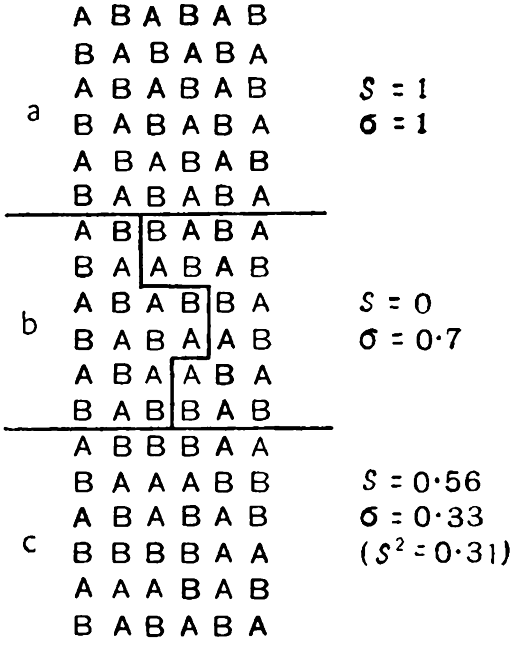


FIG. 214. The relationship between short- and long-range order.

A partially ordered solid solution consisting of a large number of antiphase regions has an increased energy on account of the surface energy of its boundaries. Raising the temperature encourages these boundaries to appear since they increase the entropy of the alloy. As the temperature is lowered these boundaries become unstable and the growth of some regions takes place at the expense of others. The final equilibrium state is reached when these regions reach the size of the individual crystals; this corresponds to the establishment of complete long-range order. It follows from the theory that there must be a critical temperature T_k below which the boundaries between the antiphase regions are unstable. Long-range order begins to be established at this temperature. The existence of antiphase regions and their growth has been investigated by means of X-rays for the alloy Cu_3Au in which the processes take place very slowly on account of the peculiarities of the structure. The solid Cu_3Au has the face-centred cubic structure formed by the superposition of four separate simple cubic structures and containing four families of sites. Thus in the alloy Cu_3Au there are four systems of antiphase regions, the growth of which proceeds comparatively slowly even at those temperatures when the rate of diffusion is large. For this alloy $T_k = 390°$ C although at temperatures of the order of 250° C the rate of interdiffusion of atoms of gold and copper is not insignificant; the alloy is easily supercooled and as regards long-range order remains disordered.

Jones and Sykes showed by means of X-rays that there were antiphase regions in a supercooled alloy. The breadth of the boundaries was shown by X-rays and by thermal measurements to be very small, of the order of two interatomic spacings. This gives support to the hypothesis of nearest neighbour interactions. Thermal treatment must change the dimensions of the antiphase regions from about 10 atomic diameters up to the dimensions of the individual crystals. The establishment of long-range order even at temperatures a few degrees below T_k requires about 50 hours. However in the alloy CuZn which has a b.c.c. structure there are only two families of antiphase regions, so here long-range order is established at a much greater rate.

Problems

1. Calculate the number of bonds n_{AA} and n_{AB} as a function of s for alloys of the type (*a*) AuCu, (*b*) $AuCu_3$.
2. Calculate the dependence of the number of bonds n_{AB} on concentration for an alloy of the type (*a*) AuCu, (*b*) $AuCu_3$.
3. For an alloy of the CuZn type calculate the change of the Kurnakov temperature $T_k(c)$ as a function of concentration c.

Chapter 13

FORMATION AND GROWTH OF CRYSTALS

General Observations

At sufficiently low temperatures the equilibrium form of the majority of substances is the crystalline state. Phase transitions from the gaseous and liquid states to the crystalline state (crystallization) and also the processes of phase transition and recrystallization which take place in the solid state are in general of first order.

A phase, which is stable at high temperatures, becomes less stable relative to the new crystalline phase as the temperature is lowered. Let F_A be the free energy of the melt, F_B that of the crystal. The temperature of phase equilibrium (of melt and crystal) T_0 is the temperature of crystallization or the melting point. Above the temperature T_0, $F_A < F_B$, and consequently the melt is stable. Below T_0, $F_B < F_A$, and consequently the crystalline phase is the more stable. The condition $F_B < F_A$ determines the thermodynamic requirement for the process of crystallization to occur but says nothing about its mechanism and kinetics. In the first instance we have to rely on experiment to give an answer to all those problems which are important in the theory and practice of production of crystalline substances. Generalisation of the experimental results will then lead to theoretical understanding of the stages of crystallization.

1. Basic Experimental Results Common to many Crystals

The Rate of Formation of Crystallization Nuclei

It is observed that when the melt is cooled to the temperature of crystallization or to some lower temperature, the whole mass of the melt does not transform to the crystalline state instantaneously nor simultaneously throughout its whole volume. Crystallization proceeds locally from one or more centres with a finite velocity taking in the whole volume of the unstable phase progressively. The velocity c at which the centres of crystallization on nuclei are formed is equal to the number of nuclei formed in unit volume of the alloy in unit time. c has dimensions $[L^{-3}\,T^{-1}]$. The nuclei of crystallization may be particles of the same crystalline substance or particles with similar composition and structure. The process of seeding is sometimes carried out in order to obtain crystals of required magnitude, shape, and orientation. Nuclei of crystallization can also arise as a result of fluctuations. The velocity at which these arise is strongly dependent on the degree of supercooling ΔT of the alloy or the degree of supersaturation Δp of the vapour or solution.

The Table reproduced below is due to Tamman who measured the velocity at which crystallization nuclei were formed in supercooled betol (an organic compound with melting point $T_0 = 95°$ C).

Velocity of Formation of Crystallization Nuclei in Betol (in mm^{-3} per minute)

$T°$, C	35	30	20	16	12	8	0	−10
c	0	1	8	40	28	4	2	0

The velocity of spontaneous formation of nuclei increases rapidly as the degree of supersaturation increases but with high supersaturation begins to fall again. This is explained by the fact that when the temperature of a condensed phase is lowered the viscosity rises rapidly, the mobility of the atoms falls, and the rate of diffusion processes is reduced.

Temperature Dependence of the Viscosity of Betol

$T°$, C	50	40	30	20	10	0
η, poise	0·014	0·036	0·157	2·65	46·0	580

Because of these factors the curve $c = c\,(\Delta T)$ has a maximum (Fig. 215).

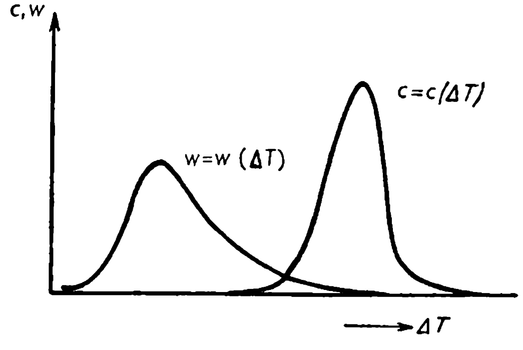


FIG. 215
Curves of the rate (c) of formation of crystallization nuclei and the linear velocity (w) of their growth.

Linear Rate of Growth

On the boundary between the crystal and the melt there is continuous deposition of new parts of the crystalline phase. As a result of this, the boundary between the crystal and the melt moves into the melt at a rate which is determined by the degree of supercooling. The linear rate of growth w is the second characteristic parameter of the process of crystallization.

Linear Rate of Growth of Crystals of Betol

$T°$, C	80	78	75	73	68	66	65
w, mm/min.	0·93	1·05	1·00	0·97	0·86	0·80	0·50

The curve $w = w\,(\Delta T)$ has also a maximum (Fig. 215). The similarity of the curves of c and w is not fortuitous. Deposition of crystalline material on the surface of the growing crystal proceeds discontinuously this being connected with the formation of two-dimensional nuclei on the crystal faces. When a two-dimensional nucleus has been formed, deposition of a new crystalline layer (growth of the two-dimensional nucleus) proceeds all over the face of the crystal at a comparatively large rate. The temperature dependence of the rate of formation of two-dimensional nuclei is in general similar to the rate of formation of three-dimensional nuclei. At small degrees of supercooling both c and w increase because of the increase in the difference between the free energies of the melt and the crystal, while at larger degrees of supercooling the curves of c and w both fall because of an increase in viscosity.

The Glassy or Amorphous State

When a liquid is cooled there is not in every case a transition to a crystalline state. A variety of substances have a strong tendency to supercooling and the ability to remain for long periods in the supercooled state. At sufficiently low temperatures, these supercooled liquids exhibit mechanical and thermal properties which are characteristic of solids; their high viscosity and the small diffusional mobility of the atoms, lead to the retention of a definite spatial distribution of atoms. Such a body is elastic.

Thermal agitation consists of vibrations of atoms around fixed positions of equilibrium. However, in distinction to crystals the centres of vibration of the atoms in a liquid do not form a regular space lattice. There is no long-range order in the distribution of the atoms in such systems—only short-range order of the type we associate with liquids. A typical example of a substance which has a marked tendency to supercooling is glass, so the corresponding state of a substance with large viscosity, elasticity and a disordered distribution of atoms is often called *glassy* or *amorphous*. In distinction to the process of crystallization which takes place at a definite temperature T_0 the transition to the glassy state with high viscosity takes place over a more or less wide range of temperature.

From the phenomenological point of view the difference between substances which readily crystallize and substances which are readily supercooled is explained by differences in the relative positions of the maxima of the curves of $c = c(\Delta T)$ and $w = w(\Delta T)$, since on the whole the process of crystallization depends on the rate of formation of nuclei as well as on their linear rate of growth. In the case where the maxima of both curves are coincident or close together, the substance crystallizes easily and it is difficult or impossible to bring it into an amorphous state by cooling the melt. Metals are such substances. In the case where the maxima are not superposed and the curves do not cut each other the substance readily passes into the glassy state, and in distinction to metals, it is difficult to crystallize it (glass). When the curves of c and w are partially coincident the substance may be brought either into the amorphous or into the crystalline state depending on the rate of cooling. An example is selenium.

According to crystal chemistry, the reason that glass has such a pronounced tendency to pass into the amorphous state is that the basic structural motif in glass is a net of silicon-oxygen tetrahedra, each linked with the other by directed valence bonds composed of oxygen atoms. The formation of a regular periodic structure in such conditions is very difficult. Not only glass but also quartz—that is, pure silicon dioxide—can easily be produced in amorphous form (fused quartz).

We can distinguish a variety of substances by their tendency to crystallize. Readily crystallizable substances are those metals which have structures corresponding to close-packing or nearly close-packing of spheres. Metals with more complex structures (zinc, tin, antimony, selenium) have a greater tendency to supercooling. Boric anhydride, which is built up of boron-oxygen triangles and tetrahedra, has a particularly marked tendency to supercooling. The formation of quite small crystalline aggregates of this substance on the glass walls of a vessel takes about a week, while to crystallize the whole mass requires several months. Many molecular organic substances also readily form glasses. The reason suggested by crystal chemistry for this is that the formation of a regular crystalline structure in these cases requires not only periodic distribution of the centres of the molecules but also a strictly determined relative orientation of the molecules in space corresponding to their orientation in the crystal.

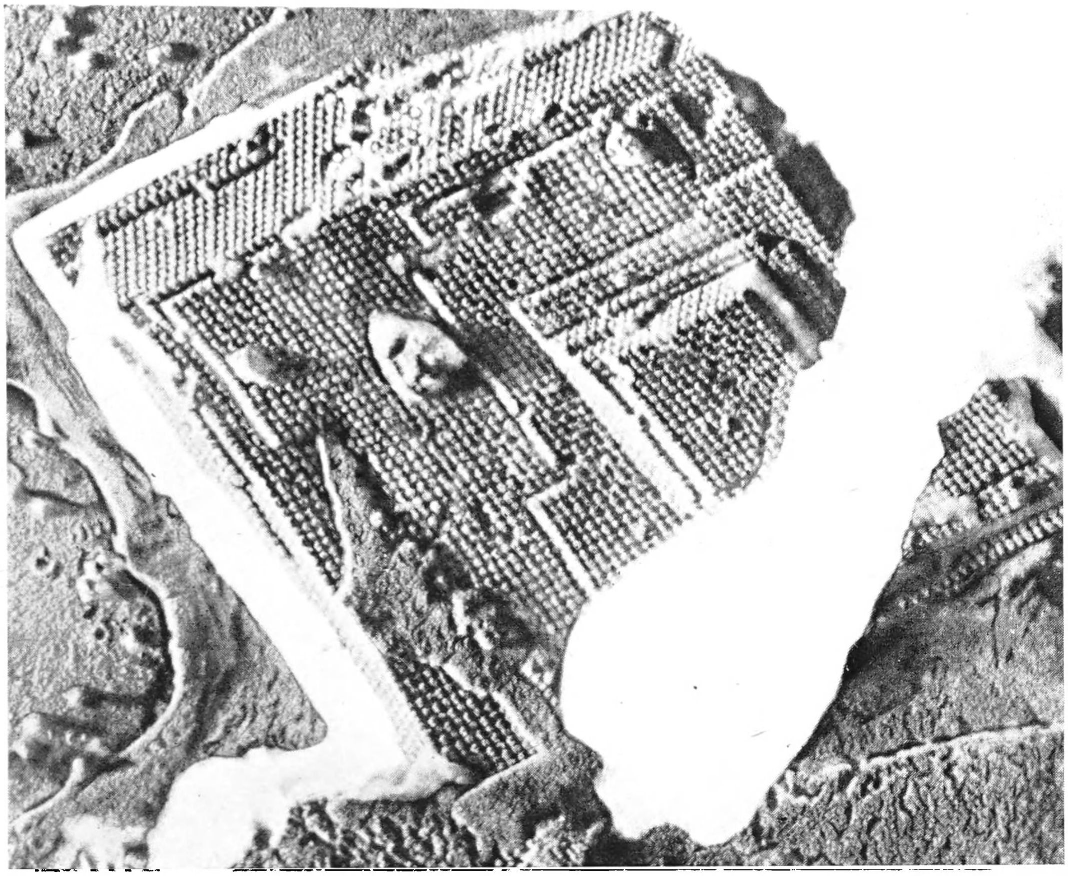
Terraces and Spirals on Crystal Faces

Studies of the shapes of crystals show that real crystals are never bounded by completely plane faces. Instead, various types of departure from plane faces are encountered, and these have a great influence on a variety of crystal properties. Investigation of the dissolution of faces makes it possible to pick out the basic lines of facial sculpture and discover the important properties of growth.

The most distinctively sculptured faces are those which appear on crystals formed from the vapour or from solution. A feature of such crystals is steps and terraces (Fig. 216), which show that growth takes place by a layer process and is connected with the formation of surface nuclei. When a growing layer reaches the edges of the crystal its growth stops. Its growth may also be stopped by obstacles such as are met with on crystal surfaces. Lemlein produced interesting shapes of growth fronts, in particular the bending of growth fronts round obstacles and the production of curved growth fronts by allowing a plane front to grow through a small aperture (similar to the phenomenon of diffraction and the production of elementary Huyghens' waves).

Comparatively recently an interesting detail of the sculpture of a crystal face was discovered, namely the spiral. Fig. 217*a* shows circular spirals on the face of a silicon carbide crystal. More frequently, the spirals have polygonal shapes corresponding to the symmetry of the faces. Fig. 217*b* shows a hexagonal spiral again forming on a silicon carbide crystal. Fig. 217*c* shows a rhombic spiral on a paraffin crystal and Fig. 217*d* shows a two-start rectangular spiral appearing on the face of a crystal of potassium ferrocyanide (see also Fig. 228).

Step-height measurements on the spirals have been carried out, especially on silicon carbide, by a micro-interference method, and have shown that the heights are equal to or multiples of the height of the unit cell. The spiral form of the faces makes it possible for a crystal to grow without interruption since, when the plane fronts spread out, the spiral rotates and thus the stepped layer never grows out. Spiral growth of crystals explains several curious structural properties of complex crystals such as the multitude of structural types (polytypes) and the appearance of large periods (superperiodicity) which are found in silicon carbide, zinc sulphide, iceland spar, and so on. Surface spirals are found on a wide range of substances with various chemical compositions and a variety of physical properties (silicon carbide, cadmium iodide, beryllium, silver, paraffin and others). This shows that spirals are not an accidental phenomenon on crystals, but are one of the typical features of crystal sculpture and an important mechanism whereby crystals grow from the vapour and from solution.



Photograph by R. W. G. Wyckoff

FIG. 216. Stepped shape of crystal faces (compare Fig. 66).

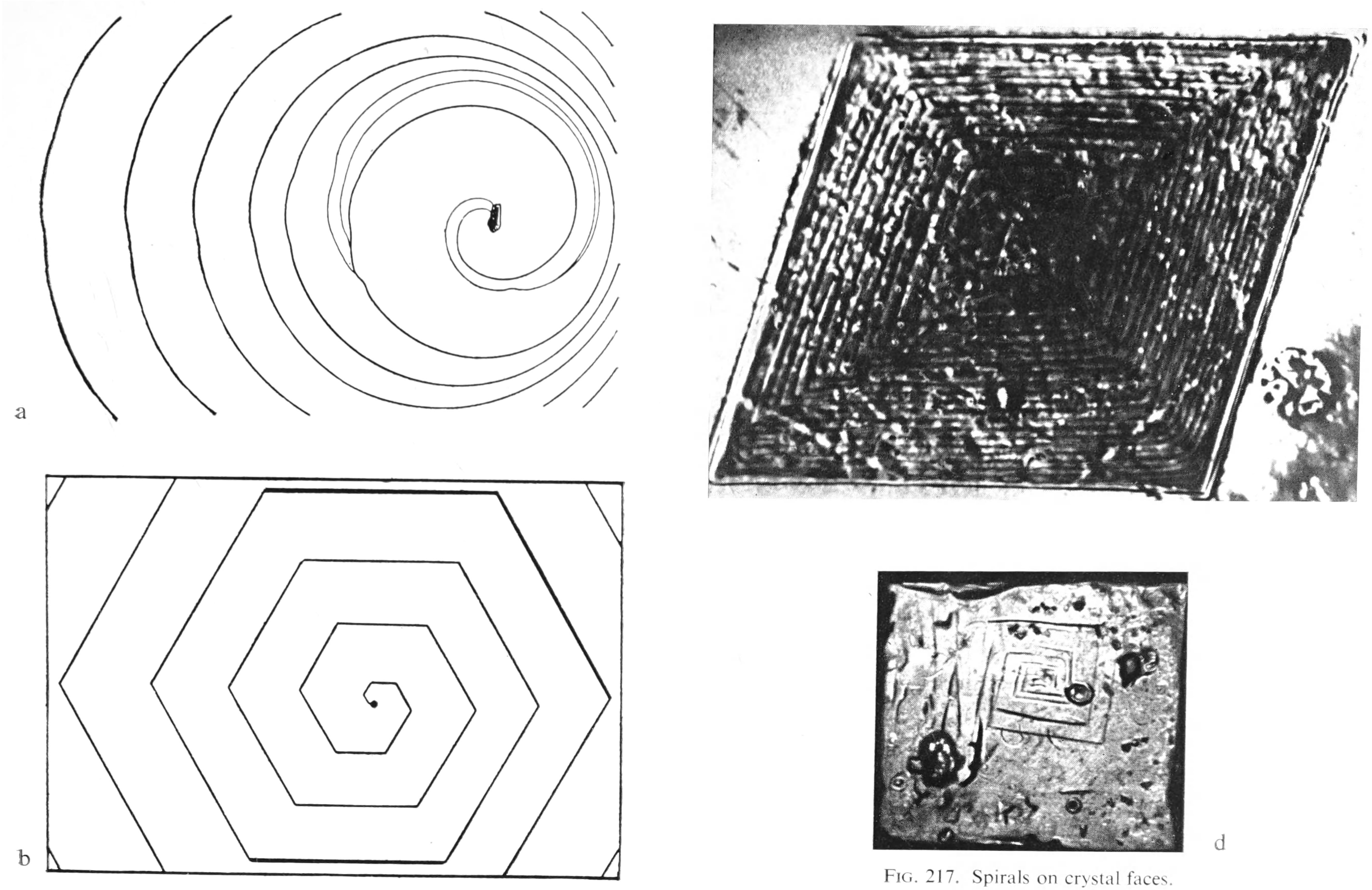


FIG. 217. Spirals on crystal faces.

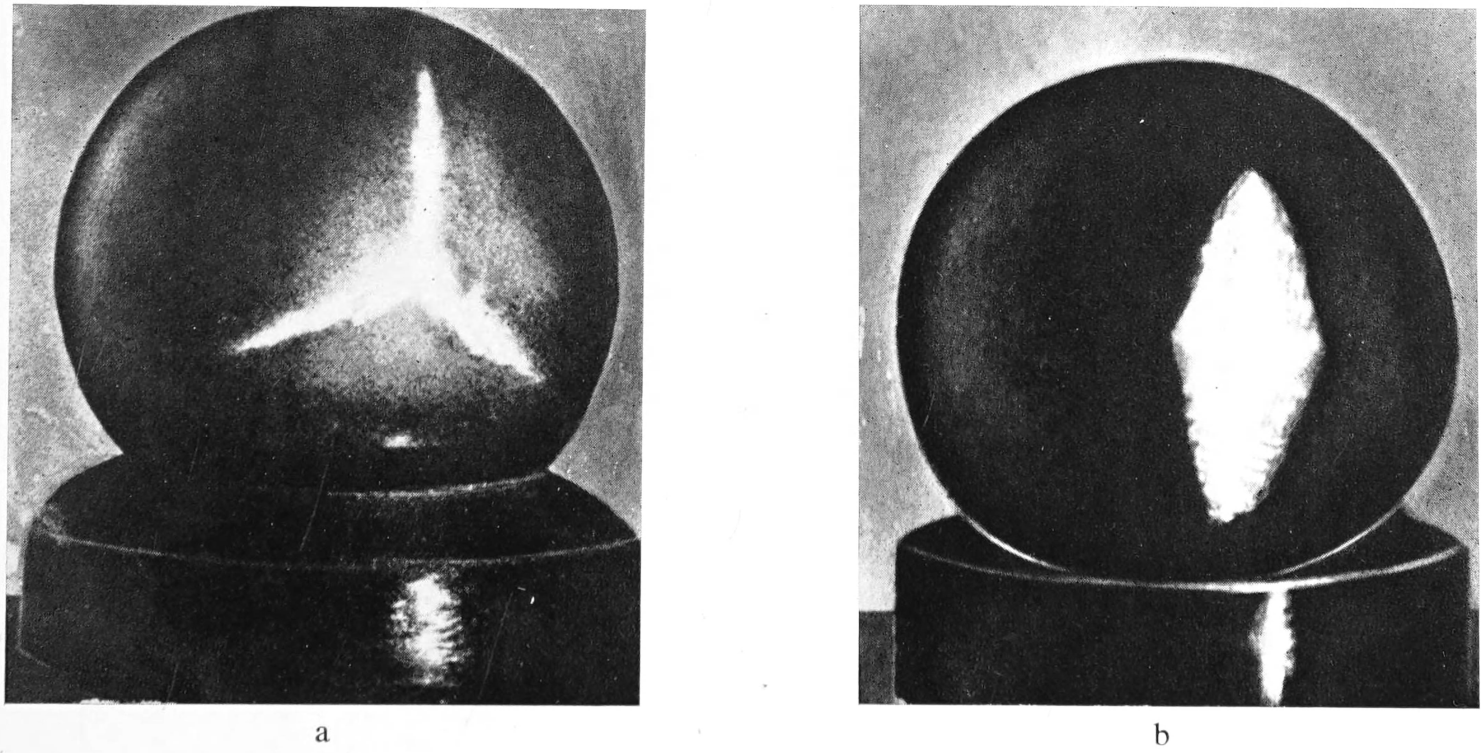

a
b

FIG. 218. Reflexion figures on the surface of a rock salt sphere in the directions (a)—[111]; (b)—[110].

Surface Diffusion of Atoms

The rate of migration of atoms on the surface of a solid (surface diffusion) is significantly greater than the rate of diffusion of atoms within a solid. This depends on the fact that the activation energy for surface diffusion is several times less than the activation energy for bulk diffusion. There is direct experimental evidence for high atomic mobility on crystal surfaces; thus Volmer, studying the growth of plate-shaped crystals of mercury from the vapour, showed that although the majority of vapour atoms condensed over the whole area of the face and not on its perimeter, growth of the crystal did not proceed in such a way as to thicken the plate but in the direction of increasing area of the platelet. This can be explained if it is assumed that atoms adsorbed on the surface of a plate-shaped crystal migrate forthwith to the edges and that growth of the crystal proceeds there. A direct experiment to show the effect of surface diffusion was carried out by Lukirsky, who produced a spherical shaped crystal of rock salt; after heating for several hours at temperatures of between 720-760° C, that is well below the melting point, figures having the symmetry of the crystal were observed on the spherical surface. Around the places where the [100], [110], and [111] axes emerged, there were formed respectively square, rhombic and triangular figures (Fig. 218). The formation of these figures can be explained by supposing that the heating gave rise to surface diffusion which caused the crystal to move closer to a more equilibrium form and, from being a sphere, become transformed into a polyhedron with stepped faces. The departure from spherical shape was also supported by the observation that the annealed crystal did not roll along a straight line path but along a zig-zag trajectory. Changes in the linear dimensions of the crystal as a result of surface diffusion were very small, being about 10 microns for a sphere of diameter 2 cm.

Total Rate of Crystallization

Consider a system existing under given conditions of supercooling. Such a system might be a melt or an unstable crystalline phase (the latter case includes phase transitions or the recrystallization of a deformed metal). At the start the amount of material transforming is zero. With the passage of time this quantity increases but the whole mass of the unstable phase is not transformed into the stable phase. In practice it is important to know the total transition rate, that is the rate of increase of the amount of new phase at each instant and also the time of complete transition t_r.

In Fig. 219 the unbroken curves show the kinetics of phase transition of γ- into α-iron (based on magnetic measurements). The temperature of α-γ equilibrium in iron

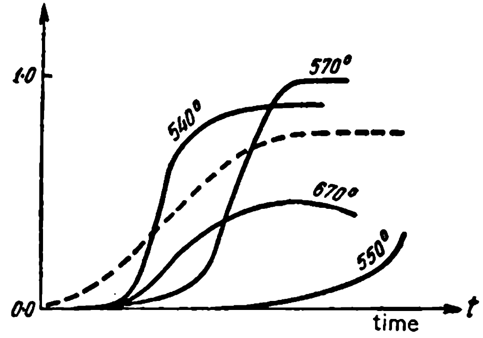


FIG. 219. Total rate of crystallization.

is 906° C. and the transition proceeds at a rate which depends on the degree of supercooling. At the beginning the transition rate is small because of the small number of nuclei and the short time available for their growth. As time goes on the rate of transition increases steeply and dies away equally quickly towards the end of the transition. Similar S-shaped curves occur in the transition of rhombic sulphur into monoclinic (dotted curve, Fig. 219). The decrease in the transition curve in the final stages is explained by collisions of the growing crystals with each other and with the walls of the vessel, resulting in a cessation of growth on the common faces.

2. Thermodynamics of Crystal Formation

Difference of Free Energies of Crystal and Melt

Consider a system consisting of a crystal and its melt. A change in the specific volume free energy of these phases as a result of a small degree of supercooling $\Delta T = T_0 - T$ can be written in the form

$$\Delta F_A = S_A^0 \Delta T \quad \text{and} \quad \Delta F_B = S_B^0 \Delta T,$$

where S_A^0 and S_B^0 are the entropies of the melt and the crystal at the temperature of crystallization. The difference between the free energies of these phases is

$$\Delta F' = (S_B^0 - S_A^0)\Delta T = -\frac{q}{T_0}\Delta T, \qquad \ldots [13.1]$$

where q is the latent heat of crystallization. Formula [*1*] takes account of only the volume free energy and not of the surface energy which can be neglected for a sufficiently large crystal. If, in the supercooled melt there is a crystal such that its free energy is less than the free energy of the melt, then the crystal will begin to grow and will continue to do so for a time but the whole melt will not crystallize. For this to happen it is necessary that sufficient heat be removed to maintain the required degree of supercooling.

The equilibrium condition for a sufficiently small crystal is quite different because of the influence of the surface energy. For simplicity suppose that the crystallite has the shape of a sphere of radius r. Taking account of the surface energy and the dimensions of the crystal, formula [*1*] takes the form

$$\Delta F = -\tfrac{4}{3}\pi r^3 \left(\frac{q\Delta T}{T_0}\right) + 4\pi r^2 \sigma, \qquad \ldots [13.2]$$

where σ is the surface tension at the crystal-melt interface.

Heat of Formation of Nuclei of Crystallization

The change in the free energy difference between crystal and melt is shown as a function of the size of the crystal in Fig. 220. For a very small crystallite, the balance of free energy is seen to be positive on account of the second term and, consequently, such a crystallite will be unstable relative to the melt. For a crystallite to be more stable than the melt, its dimensions must be greater than some critical value r_k corresponding to the maximum of the curve of ΔF. All crystallites with dimensions less than the critical will remelt while those with dimensions greater than critical become centres or nuclei of crystallization. Since the melt is unstable relative to them such crystallites will grow. The critical dimension of a crystallization nucleus is obtained from the condition

$$r_k = \frac{2\sigma T_0}{q(T_0 - T)}. \qquad \ldots [13.3]$$

The greater is the latent heat of crystallization and the degree of supercooling and the less the surface energy, the less is the critical size of a nucleus.

The heat of formation of a nucleus is obtained from the height of the maximum of the free energy curve (Fig. 220)

$$\Delta F_k = -\frac{8\pi\sigma}{3} r^2 + 4\pi\sigma r^2 = \tfrac{4}{3}\pi\sigma r^2 = \tfrac{1}{3}\sigma S, \qquad \ldots [13.4]$$

where S is the surface area of the nucleus· The heat of formation is 1/3 of the surface energy of the nucleus.

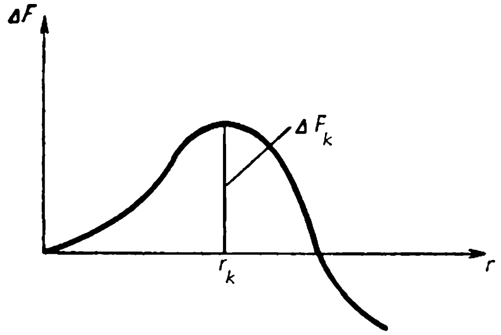


FIG. 220. Heat of formation of a crystallization nucleus.

Influence of Walls and Impurities

The energy required to form a nucleus of a new phase is significantly reduced when the nuclei are formed on bounding surfaces, the walls of the vessel or the surfaces of impurity particles. Consider the formation of a spherical nucleus ABC (phase II) on a plane surface (phase III) and bounded by the mother liquor (phase I) which feeds the crystallization nucleus (Fig. 221). In a case where the nucleus is formed on the wall, the work to form it is

$$R = \tfrac{1}{3}\sigma S = \tfrac{1}{3}(\sigma_1 S_1 + \sigma_3 S_2 - \sigma_2 S_2),$$

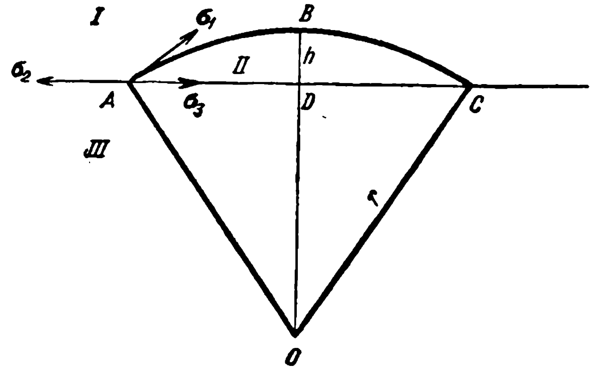


FIG. 221. Formation of a nucleus on a boundary between phases.

where σ_1 is the surface tension on the nucleus mother-phase interface I, σ_2 the same on the wall mother-phase interface, and σ_3 the same on the nucleus-wall interface. In equilibrium $\sigma_2 - \sigma_3 = \sigma_1 \cos\theta$. The area of the surface is S_1 and the area of the base of the spherical segment of the nucleus are given by

$$S_1 = 2\pi r h = 2\pi r^2 (1 - \cos\theta),$$
$$S_2 = \pi r^2 \sin^2\theta.$$

Whence

$$R = \tfrac{1}{3}\pi r^2 \sigma_1 [2(1 - \cos\theta) - \sin^2\theta \cos\theta]. \qquad \ldots [13.5]$$

For the limiting angle $\theta = \pi$ (no wetting) this expression gives the heat of formation of a free nucleus inside a homogeneous volume of the mother-phase; $R = 4\pi r^2 \sigma$. In the case of partial wetting $\theta < \pi$ and $R < \Delta F_k$ that is, the heat of formation of a nucleus on an interface is less than the heat of formation of a free nucleus. With complete wetting $\theta = 0$ and $R = 0$. In this case the formation of a new phase (condensation of the vapour) proceeds even at very small supersaturation.

These discussions are valid in general and explain the important part played by the walls and particles of impurity in the formation of nuclei. The heat of formation of nuclei in general is less for particles of isomorphic impurity or crystalline particles where one or other of the faces has a structure compatible with the crystallizing substance.

The Probability of Formation of Nuclei by Fluctuations

The process of crystallization can begin in two ways. Firstly, there may be intentional or accidental introduction of ready made centres of crystallization (crystal seeds or impurities); alternatively the nuclei may form spontaneously, the energy required being supplied by thermal fluctuations. The probability that a nucleus will be so formed is determined by the expression

$$e^{-\frac{\Delta F}{kT}} = e^{-\frac{\sigma S}{3kT}} = e^{-\frac{16\pi\sigma^3 T_0^2}{3kq^2 (T_0 - \Delta T)\Delta T^2}}. \qquad \ldots [13.6]$$

For crystallization in condensed systems (liquid or solid phase) an important factor is the increase in the viscosity as the temperature is lowered. This arises from a reduction in diffusion, the temperature dependence of which is expressed by a formula analogous to [*6*]: $D = Ae^{-Q/kT}$, where Q is the energy of activation for diffusion. Taking into account the viscosity of the system, the probability of formation of a nucleus is determined by the formula

$$c \simeq e^{-\frac{1}{kT}\left\{Q + \frac{1}{3}\pi\sigma\left[\frac{4\sigma T_0}{q(T - T_0)}\right]^2\right\}}. \qquad \ldots [13.7]$$

The turning value of the function is expressed by the equation

$$ax^2 (3-x) = Q(x-1)^3,$$

where

$$x = \frac{T_0}{T} \quad \text{and} \quad a = \tfrac{1}{3}\pi\sigma\left(\frac{4\sigma}{q}\right)^2.$$

Mechanism of Crystal Growth

Consider a very simple model of a crystal, one made up of cubic blocks, and take account only of interactions between nearest neighbours. At sufficiently low temperatures the equilibrium state of the crystal is determined by the minimum of potential energy of interaction which corresponds to the densest packing of the blocks. The formation of a step on the external face arises from there being insufficient blocks to make up a complete face (Fig. 222). At higher temperatures the crystal is in dynamic equilibrium with its vapour which has a finite density of blocks. As a result of the balance between the opposing processes of condensation and evaporation, the structure of the external faces of the crystal will be similar to that shown in Fig. 223, being permanently 'rough'.

The energies of crystal blocks are proportional to the number of neighbours and differ according to their positions. The weakest binding is for isolated blocks of type B. When the blocks congregate in groups the binding energy increases. On an incomplete face, the strongest binding is that of blocks of type A which land in a re-entrant step. When the crystal is evaporating, the first to go are isolated blocks and individual groups. The most difficult blocks to disperse are those in the complete part of the face, although a certain number of vacancies of type C will be formed by thermal energy fluctuations.

Blocks on the faces of a crystal have high mobility (surface diffusion). Thus growth of a crystal from the gaseous phase (or from solution) at given supersaturation Δp proceeds as follows: a steady stream of blocks lands on the external face of a crystal, falling in the main in positions B. These blocks move around over the surface of the crystal remaining there for some mean lifetime. If in this time the blocks move into the corner of a step then they may stick there, causing the step to grow longer and spread across the face. Thus the face of a crystal grows by means of terraces and the spreading out of plane growth layers. If as a result of growth and the filling up of steps a complete crystal face is formed then the process of growth is halted since individual blocks adsorbed on such a face will be in unstable equilibrium there unless the degree of supersaturation is changed. If the growth of the crystal is to continue, new steps must be formed and this requires the formation of stable groups of blocks on the face. These are the two-dimensional nuclei (Fig. 224).

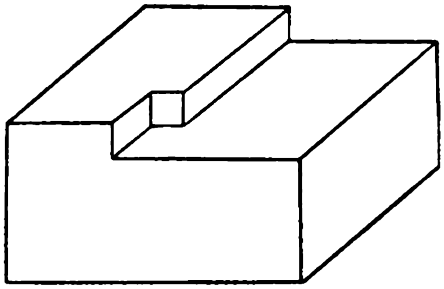


FIG. 222. Growth sequence of a crystal.

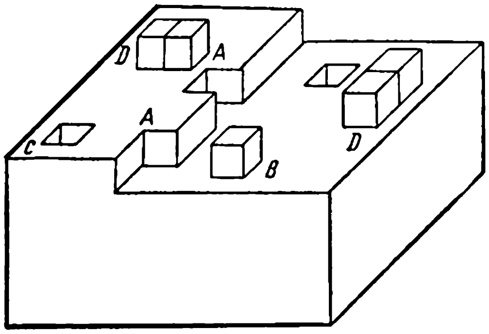


FIG. 223. Growth of a real crystal face.

The Probability of Formation of Two-dimensional Nuclei

Suppose that the nucleus is an island with the shape of a circle of radius ρ. The change of free energy required for the formation of such a nucleus is

$$\Delta F = -\pi\rho^2\sigma + 2\pi\rho\lambda, \qquad \ldots [13.8]$$

where σ is the gain of free energy per unit area when molecules are added to the island, and λ is the free energy per unit length of a step. As in the case of formation of three-dimensional nuclei, the formation of two-dimensional nuclei is connected with the work which goes to form the perimeter. From the condition for a minimum

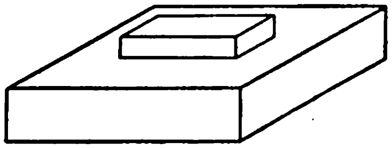


FIG. 224. Two-dimensional nucleus on the face of a crystal.

of ΔF we obtain the critical dimension and the work to form a two-dimensional nucleus

$$\rho_k = \frac{\lambda}{\sigma} \quad \text{and} \quad \Delta F = \frac{\pi\lambda^2}{\sigma} = \pi\lambda\rho. \qquad \ldots [13.9]$$

The heat of formation of such a nucleus is half the energy of the step which marks its perimeter.

The calculation of the probability of formation of two-dimensional nuclei shows strong dependence of the probability on the degree of supersaturation. If a crystal is to grow at a rate which is measurable in the laboratory (the order of 1μ per month) the degree of supersaturation required has to be not less than 25-50%. With higher degrees of supersaturation the process of growth is not held up by the formation of two-dimensional nuclei. At lower degrees of supersaturation the probability of formation of nuclei is much too small. Experiments on the growth of crystals from the vapour show that growth proceeds at a measurable rate at much lower degrees of supersaturation (of the order of 1%). The theory of growth of the ideal crystal is thus in serious disagreement with experiment.

From the above we may conclude that the experimental fact that crystals grow at small degrees of supersaturation (and in these conditions are completely regular and well-formed) shows that real crystals are not perfect. Frank and others have developed a theory of the mechanism of growth in which a step on the crystal face does not grow out but is available indefinitely. This mechanism of growth is connected with the emergence of screw dislocations on crystal faces.

Screw Dislocations and the Growth of Spirals

If one part of a crystal slips by a small amount relative to another part (Fig. 225) a step *AB* may be formed which finishes at the point *B* within the crystal face. Fig. 226 shows a model of such a defect, called a screw dislocation, on an atomic scale. When such a defect is present the face ceases to be an ideal plane becoming bent into a section of a spiral. Depending on the direction rotation we can have either right or left handed screw dislocations. When the displacements are so small that the height of the steps are equal to one or several unit cells, the departure of the crystal face from being planar is very small and special methods are required to make the defects visible.

Given that there is a screw dislocation in the crystal face, the step can advance continuously and is not eliminated by the process of crystal growth. This explains crystal growth at small supersaturations. Fig. 227 shows successive positions of a step 1, 2, 3, 4, 5 and formation of a growth spiral (*b* and *c*). To begin with, the step advances parallel to itself but, since the rate of advance of the step close to the axis of the screw dislocation is large, the steps become curved and spirals are formed. However, the curved parts of the steps have lower stability so the rate of frontal growth is here reduced until equilibrium is reached, corresponding to a spiral with a shape such that all its parts advance with the same angular velocity.

Circular spirals form when the rate of growth is independent of crystallographic direction. The rate of advance of a plane front with a radius of curvature ρ is determined by the formula

$$v = v_\infty (1-\rho/\rho_k), \qquad \ldots [13.10]$$

where v_∞ is the rate of growth of a straight front and ρ_k is the critical radius of a curved two-dimensional nucleus given by

$$\rho_k = \frac{a\Phi}{2kT\ln\alpha}. \qquad \ldots [13.11]$$

Here $\alpha = \dfrac{p}{p_0}$ is the degree of supersaturation, $\frac{1}{2}\Phi = \sigma a$ is the binding energy of nearest neighbours, $\sigma = \alpha - 1$ and a is the interatomic spacing. In these conditions the shape of the spiral is similar to the Archimedean form,

$$r = 2\rho_k\theta, \qquad \ldots [13.12]$$

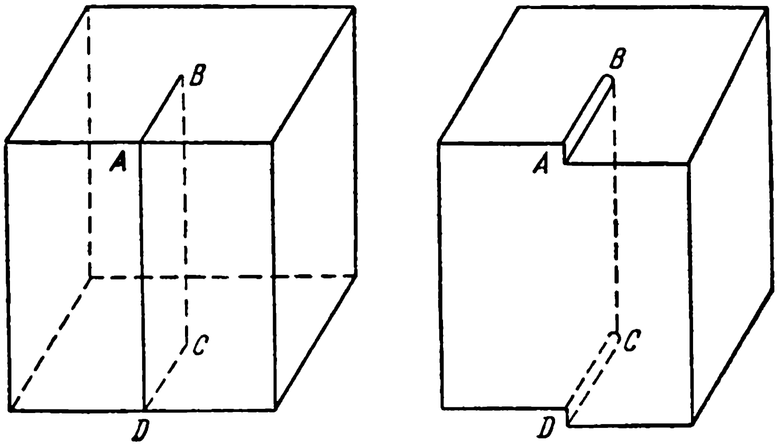


FIG. 225. Formation of a screw dislocation.

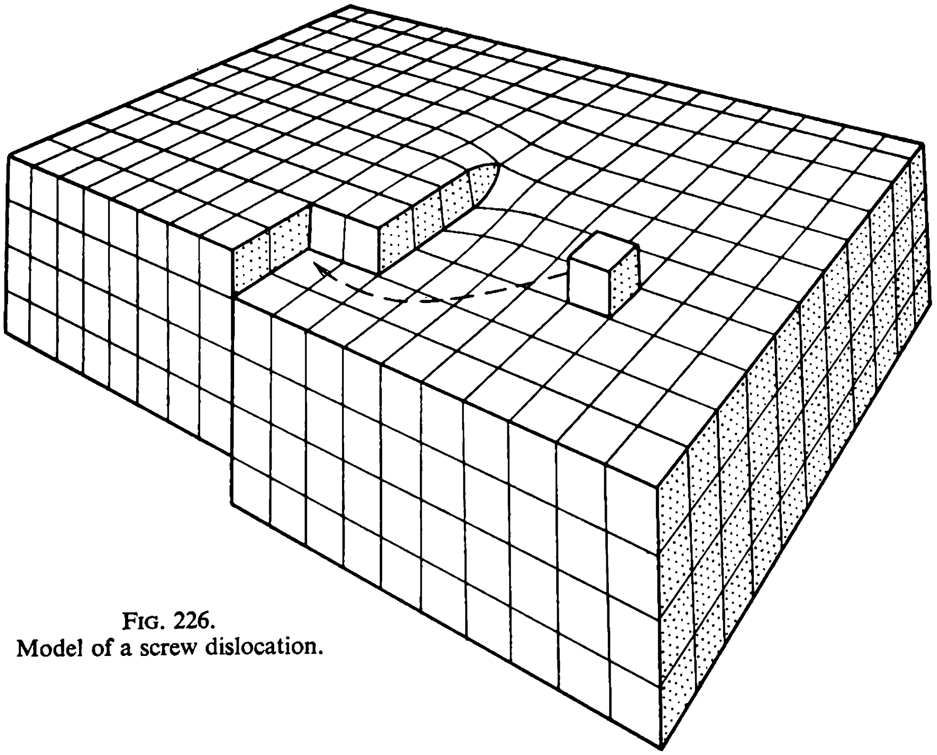


FIG. 226.
Model of a screw dislocation.

with spacing between successive turns

$$\Delta r = 4\pi\rho_k, \qquad \ldots [13.13]$$

and turning with angular velocity

$$\omega = \frac{v_\infty}{2\rho_k}. \qquad \ldots [13.14]$$

Polygonal spirals are formed when the rate of movement of a growth front v_∞ depends on crystallographic direction. Then the lines of the growth fronts tend preferentially to be parallel to the directions of dense packing on the crystal face. More complex facial relief is formed when there is interaction between several spirals with the same or opposite hands. Fig. 228*a* shows the superposition of a large number of spirals with the same direction of turn (the number of spirals is about 10^4 per cm^2) and Fig. 228*b* shows the pattern formed by the interaction of two spirals rotating in opposite directions.

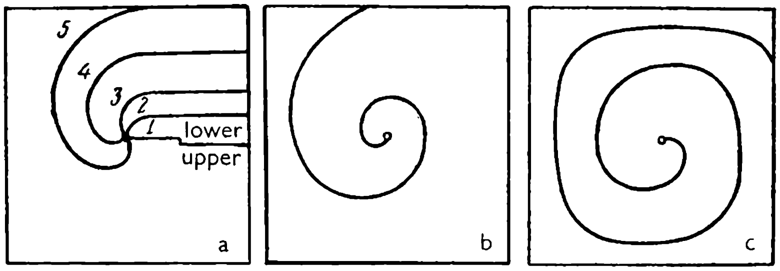


FIG. 227. Formation of a growth spiral.

3. Kinetics of Crystallization

Consider an isothermal process of crystallization from the melt and assume that the rate of formation of nuclei *c* and the linear rate of growth *w* depend only on the supercooling ΔT.

Theory, taking no Account of Collisions between the Growing Crystals

Göler and Sachs considered the kinetics of a process of crystallization under conditions which are independent of growth. Suppose at the time τ one of the nuclei is formed. At the later time t the volume to which this crystal has grown, assuming its shape to be cubic, is $\phi = 8\, w^3\, (t-\tau)^3$. Denote by $u(\tau)$ the noncrystallizing part of unit volume of the liquid at time τ. In the time interval $d\tau$, $dN = cu(\tau)\, d\tau$ nuclei of crystallization are formed. The volume of these nuclei at time t is $cu(\tau)\, d\tau 8w^3\, (t-\tau)^3$. The total volume of all the crystals which have grown since the time when crystallization began is

$$1-u(t) = 8cw^3 \int_0^t u(\tau)(t-\tau)^3\, d\tau.$$

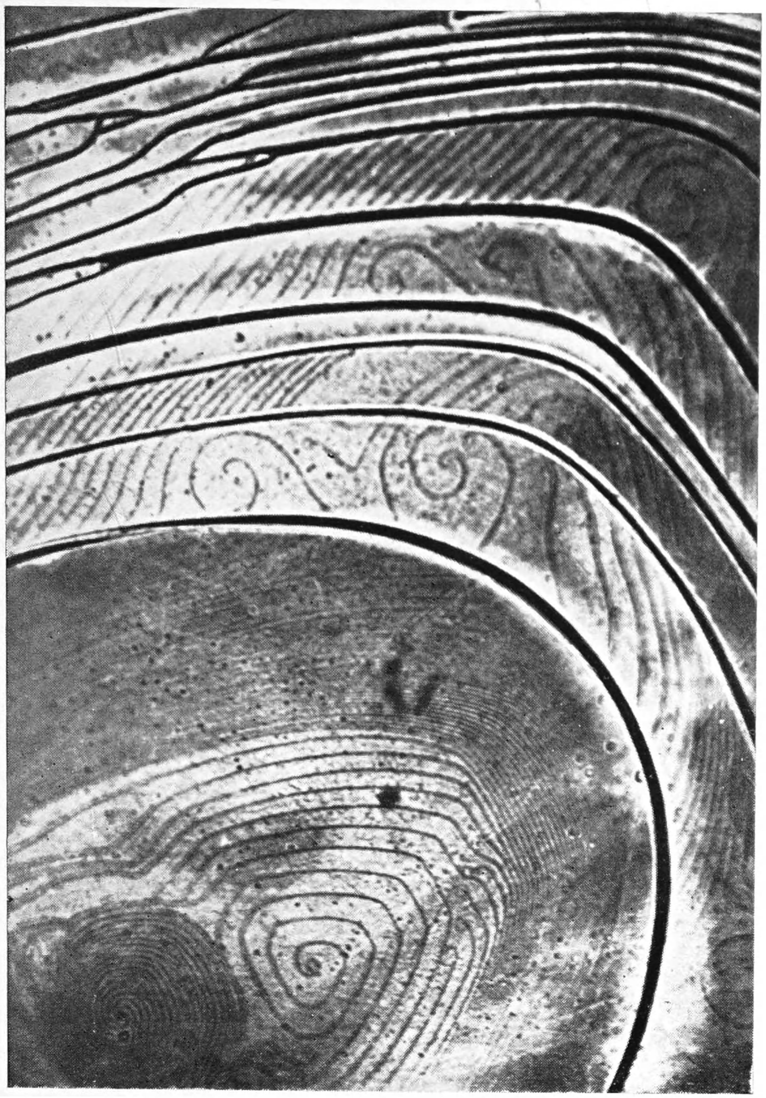
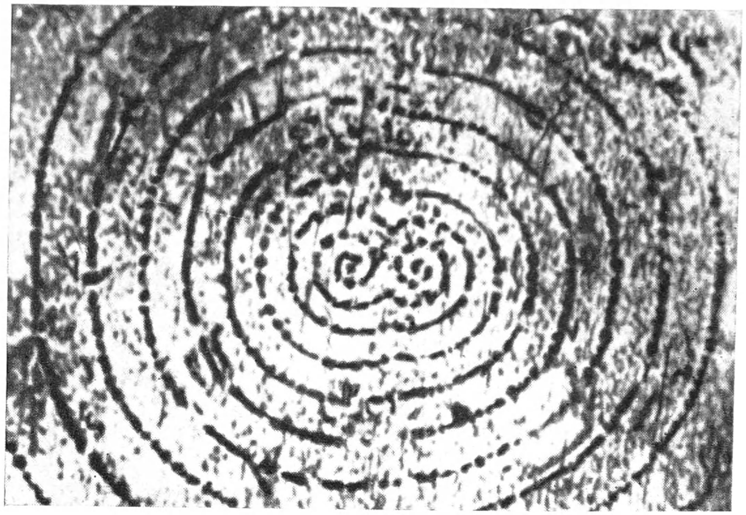


Fig. 228

(a) Spirals with the same direction of rotation; (b) superposition of two spirals with opposite directions of rotation (silicon carbide).



Hence the non-crystallizing part of the volume of the melt is

$$u(t) = 1-8cw^3 \int_0^t u(\tau)(t-\tau)^3\,d\tau. \qquad \ldots [13.15]$$

This is an integral equation of the Volterra type. Its solution is given by the series

$$u(t) = 1-\frac{4p^4t^4}{4!}+\frac{(4p^4t^4)^2}{8!}-\ldots+\frac{(-4p^4t^4)^n}{(4n)!}, \qquad \ldots [13.16]$$

where

$$p^4 = 12cw^3. \qquad \ldots [13.17]$$

This solution can be written in the form

$$u(t) = \cosh pt \cos pt. \qquad \ldots [13.18]$$

The physically interesting range of the solution lies within the limits of $\frac{1}{4}$ of a period. When $t = 0$, $u(0) = 1$; when $t = \frac{\pi}{2}$, $u\left(\frac{\pi}{2}\right) = 0$, that is the process of crystallization is completely finished. Hence the time for complete crystallization is

$$t_r = \frac{\pi}{2p} = \frac{1{\cdot}57}{p} = \frac{0{\cdot}85}{\sqrt[4]{cw^3}}. \qquad \ldots [13.19]$$

The total number of crystals in unit volume is

$$N_r = c\int_0^{t_r} u(\tau)\,d\tau = \frac{0{\cdot}05c}{p}. \qquad \ldots [13.20]$$

The maximum linear dimensions of the crystals is

$$l_m = 2wt_r = \frac{\pi w}{p} = \pi\sqrt[4]{\frac{w}{12c}}. \qquad \ldots [13.21]$$

The theoretical curve of $1-u(t)$, which represents the crystallizing part of the volume of the melt is shown on page 483. This curve rises slowly in the early moments of crystallization, rises steeply in the middle and is specially steep at the end of the process of crystallization. The initial part of the curve may agree with experiments but in the later stages of the process of crystallization there is severe disagreement since the experimental curves are S-shaped.

As we have already shown, the slowing up of the process of crystallization is explained by the fact that when the boundaries of the growing crystals meet each other or collide with each other, growth on these boundaries stops. At the end of the process of crystallization the majority of the crystals are in contact with each other and are cut off from the process of crystallization. These collisions between the growing crystals have not been taken account of in the above theory, which is thus an approximation describing only the early stages of the process when the number and dimensions of the nuclei are small and their total volume does not exceed 15-20% of the total volume.

Theory, taking Account of Collisions between the Growing Crystals

Kolmogorov has developed a theory of kinetics of crystallization for the three-dimensional case taking account of the fact that the boundaries of the growing crystals come in contact. Suppose we have a volume V of the melt (Fig. 229). Consider an moment t not fall within a volume which is crystallizing. Consider any other point P′. If at the moment $\tau < t$ a nucleus is forming at the point P' then in time $(t-\tau)$ it will have reached a distance $w(t-\tau)$ and if $w(t-\tau) > PP'$ then the point P will be engulfed by the growing crystal.

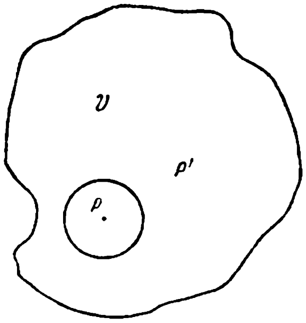


Fig. 229. Calculation of the probability of collisions of the boundaries of crystals during growth.

Consequently the critical volume around the point P is $V(\tau) = \frac{4}{3}\pi w^2 (t-\tau)^2$. Now take the unit of time to be so small that $c \ll 1$. The probability of formation of a nucleus in the volume considered is $cv\Delta\tau$. The probability that in this interval of time a nucleus will not appear within the critical volume is

$$1 - cv(\tau)\,\Delta\tau.$$

Now divide the time intervals t into units of $\Delta\tau$, $t = n\Delta\tau$, $\tau_i = i\Delta\tau$. Then the probability that the point P will not be engulfed by crystals in time t is expressed by the product of probabilities

$$q(t) = \prod_{i=1}^{n} [1 - cv(i\Delta\tau)\Delta\tau]. \qquad \ldots [13.22]$$

From which

$$\ln q(t) = \sum_{i=1}^{n} \ln [1 - cv(i\Delta\tau)\Delta\tau] = -\sum_{i=1}^{n} cv(\tau)\Delta\tau = -c\int_0^t v(\tau)\,d\tau = -c\Omega.$$

Since the non-crystallizing part of the volume is equal to the probability we have obtained

$$u(t) = q(t) = e^{-c\Omega},$$

$$\Omega = \int_0^t v(\tau)\,d\tau = 8w^3 \int_0^t (t-\tau)^3\,d\tau = 2w^3 t^4,$$

$$u(t) = e^{-2cw^3t^4} = e^{-\frac{p^4 t^4}{6}}. \qquad \ldots [13.23]$$

The first two terms of the expansion of this function in series agree with the corresponding terms of the series [*16*]. What is new is that, according to [*23*], completion of the process of crystallization requires an infinitely long time. If, however, we take for the time when crystallization is finished in practice as t_r' – the time in which 0·999 of the volume of the mother-phase has crystallized then

$$t_r' = \frac{2\cdot54}{p} = \frac{1\cdot37}{\sqrt[4]{cw^3}}. \qquad \ldots [13.24]$$

The number of crystals at the end of crystallization

$$N(t) = N(\infty) = c\int_0^\infty u(\tau)\,d\tau = c\int_0^\infty e^{-\frac{p^4t^4}{6}}\,d\tau = \frac{\sqrt[4]{6c}}{4!\,p} = \frac{0\cdot065c}{p}.$$

The value of this integral can be found from a general formula†

$$\int_0^\infty e^{-wt^n}t^z\,dt = \frac{1}{z+1}\,\frac{\frac{z+1}{n}!}{w^{\frac{z+1}{n}}}, \quad (w > 0;\ n > 0;\ z > -1).$$

Problems

1. Find the temperature of the maximum of the curve $c = c(T)$ which gives the rate of formation of nuclei.
2. Determine the critical dimension and rate of formation of a nucleus in the crystallization of (*a*) water, (*b*) copper.
3. Find the solution of the equations of the kinetics of crystallization not taking account of collisions between the growing crystals.
4. (*a*) Compare the kinetics of crystallization $\mu_{\text{I}}(t) = \cosh pt \cos pt$ and $u_{\text{II}}(t) = \exp(-p^4t^4/6)$; (*b*) find the time of complete crystallization t_r; (*c*) find the total number of crystals at the end of crystallization.
5. Find the equation for the kinetics of crystallization for the two-dimensional case taking account of collisions between the growing two-dimensional crystals.
6. Find, by the method of dimensions, the relations between the quantities which determine the kinetics of crystallization for (*a*) the three-dimensional case, (*b*) the two-dimensional case.
7. Suppose that the rate of growth of a nucleus depends on its dimensions and is proportional to the difference of free energies between the nucleus and the melt and find (*a*) the linear rate of growth w_r and (*b*) the volume rate of growth v_r.
8. Consider the kinetics of crystallization for the case when there are c_0 nuclei (per 1 cm^3 of the melt) at the beginning and in addition, new nuclei are formed at a rate c.

---

† Jahnke and Emde, *Tables of Functions*. Dover 1946.

Chapter 14

THERMAL PROPERTIES OF CRYSTALS

1. The Specific Heat of Crystals: Experimental Data

The Specific Heats c_p and c_v

The specific heat c is a measure of the change of internal energy E of a body with temperature

$$c = \frac{dE}{dT} = \phi(p, v, T). \qquad \ldots [14.1]$$

Because of the relation between p, v, and T which is expressed by the equation of state $f(p, v, T) = 0$, the specific heat depends on only two of the thermodynamic parameters and is represented graphically by a surface of some kind. The most important two plane sections of this surface are at v = constant and p = constant, and these determine the specific heat at constant volume c_v and at constant pressure c_p

$$c_v = \left(\frac{\partial E}{\partial T}\right)_v \quad \text{and} \quad c_p = \left(\frac{\partial E}{\partial T}\right)_p. \qquad \ldots [14.2]$$

c_v is the easier to calculate theoretically. However since it is difficult to keep the volume of a liquid or solid constant when it is heated, one usually measures c_p. In isobaric heating of a body the heat of expansion has to be supplied in addition and this is given by $\int p dv$. Therefore $c_p > c_v$. The difference in the specific heats is determined by the thermodynamic relation

$$c_p - c_v = \frac{\alpha^2 v T}{\kappa}, \qquad \ldots [14.3]$$

where α is the coefficient of volume of thermal expansion, κ is the coefficient of com-

Table 64

Atomic Heats of Various Metals

Metal	T, °C	c_p	c_v	T, °C	c_p	c_v	T, °C	c_p	c_v
Cu	20	5·8	5·65	500	6·7	6·2	1000	7·5	6·5
Ag	20	6·0	5·75	500	6·7	6·0	900	7·3	6·1
Au	20	6·15	5·8	500	6·7	6·0	1000	7·4	6·1
Pt	20	6·1	5·95	500	6·8	6·4	1600	7·45	6·65
Pb	20	6·35	5·9	500	7·3	6·4	—	—	—

pressibility, v is the volume of the gram molecule and T is the temperature. Table 64 gives values of the atomic heat of various metals.

As the temperature rises, the difference $c_p - c_v$ increases and reaches a value of ~ 1 cal/gram-atom. °K. At low temperatures $c_p - c_v \to 0$.

The Energy and Specific Heat of a Classical Oscillator

The particles in a solid execute vibratory motions around their positions of equilibrium. The amplitude and energy of these vibrations increase with rise in temperature. At small amplitudes the vibrations can be considered to be harmonic. The total energy of an oscillator is equal to the sum of its kinetic and potential energy $E = K + U$. The mean value of these energies taken over the period of oscillation are the same $K = U$; so consequently $E = 2K$, while for a perfect gas $E = K$. This twofold increase in the thermal energy of a crystal as compared with the energy of a monatomic gas can be considered as due to the doubling of the number of degrees of freedom. On the classical theory, every degree of freedom receives the same thermal energy $\frac{1}{2}kT$ and this is independent of the chemical nature and mass of the oscillator. The total energy of a three dimensional oscillator is thus $3kT$ and the specific heat $3k$. The total thermal energy of a crystal containing N atoms at temperature T will then be

$$E = 6 \cdot \tfrac{1}{2}kT \cdot N = 3kNT, \qquad \ldots [14.4]$$

from which we obtain for the specific heat of the crystal

$$c_v = \frac{dE}{dT} = 3kN. \qquad \ldots [14.5]$$

The classical theory predicts that the specific heat of crystals, like perfect gases, will be independent of temperature. The numerical value of the specific heat of a one component crystal referred to the gram-atom is

$$c_v = 3R = 6 \text{ cal/gram-atom. °K} \qquad \ldots [14.6]$$

and for crystals of compounds referred to the gram-molecule

$$c_v = 3R \cdot n = 6n \text{ cal/mole. °K} \qquad \ldots [14.7]$$

where n is the total number of particles in the chemical formula of the compound.

The Laws of Dulong and Petit and of Neumann and Kopp

Long before kinetic theory was known, Dulong and Petit (1819) pointed out that the atomic heat of the majority of metals was close to 6·2 cal/gram-atom. °K (Table 64). This agrees with our theoretical values. Table 65 gives the values of the molar heat and the corresponding shares of the specific heat per atom in various compounds.

Table 65

Molar Heats of Some Compounds at Room Temperature

Crystal compound	$NaCl$	CuS	$AgCl$	Ag_2S	$CuCl_2$	$PbCl_2$
Molar heat	12·3	11·9	13·1	18·6	18·6	18·5
Molar heat per atom	6·2	6·0	6·6	6·2	6·2	6·2

As was shown by Neumann (1831) and Kopp (1863) the molar heats of compounds are equal to the sums of the atomic heats of the elements they contain. The molar heats of compounds with composition *AB* are close to 12 cal/mole. °K and of compounds with composition AB_2 to 18 cal/mole. °K and so on. However at high and more especially at low temperatures there are marked departures from the law of Dulong and Petit and its generalisation, the law of Neumann and Kopp. Thus, the atomic heat of the diamond crystal which is 5·15 cal/gram-atom. °K at 830° C has fallen to 1·46 cal/gram-atom. °K at room temperature.

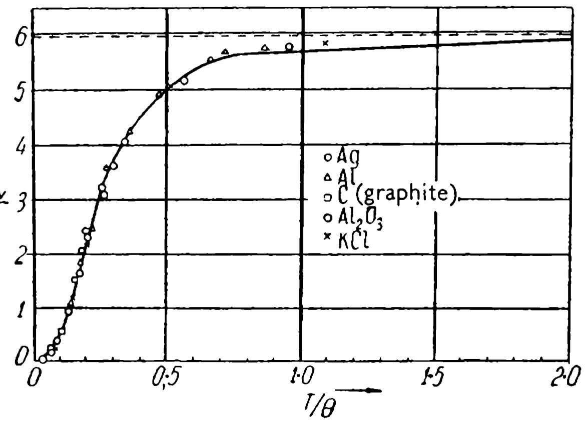


Fig. 230. Temperature dependence of the specific heat of solids.

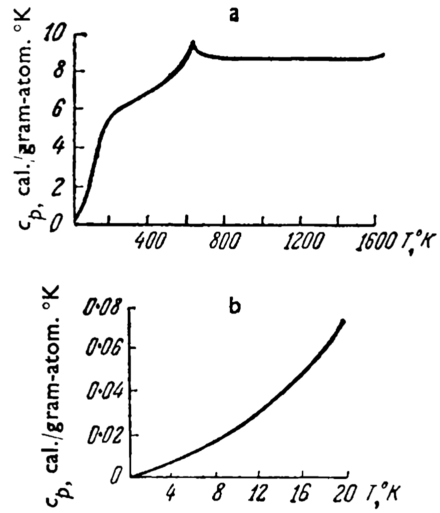


Fig. 231. (*a*) The specific heat curve for nickel; (*b*) the low temperature region of the same curve.

The Temperature Dependence of the Specific Heat

As the temperature is lowered, the specific heat of metals falls (Table 64). As the temperature approaches absolute zero the specific heat of metals and all solids approaches zero; $c_v \to 0$ as $T \to 0$ (Fig. 230). The specific heats of a variety of substances (diamond, aluminium, silver, calcium chloride, etc.) fit well on to a single curve if for each substance we measure the temperature in units of a temperature characteristic for each, and refer the specific heat in each case to a single atom. The characteristic temperature for each substance determines the upper limit of the temperature region in which there is significant departure from the law of Dulong and Petit. At low temperatures, the specific heat curve can be represented with sufficient accuracy by a T^3 law. From this the characteristic temperature can be calculated. These are given for a few substances in Table 66.

In a variety of cases there are more or less striking departures from the idealised curve of Fig. 230. At high temperatures the atomic heat of metals goes above 6 cal/mole. °K by amounts that are specially large for the transition metals. Thus the atomic heat of iron which is already 7·82 at 300°C has risen to 10·53 cal/mole. °K at 600° C. Fig. 231*a* shows how that atomic heat of nickel tends to 9 cal/mole. °K at high temperatures. At very low temperatures metals show a departure from the

Table 66

Characteristic Temperatures

Metal	θ, °K	Metal	θ, °K	Compound	θ, °K
Be	1000	Cu	315	NaCl	281
Mg	290	Ag	215	KCl	227
Zn	250	Au	170	CaF_2	474
Cd	172	Fe	420		
Hg	96	W	310		

T^3 law and change to direct dependence on T; Fig. 231*b* shows the low temperature part of the specific heat curve of nickel and the linear dependence of specific heat in the temperature range below 6°K. In this range we have for nickel

$$c_v = 0{\cdot}001744T. \qquad \ldots [14.8]$$

Linear extrapolation of this law to high temperatures shows that at 1000° K the part of the specific heat which would correspond to this law is 1·7 cal/mole. °K. This fact and also the fact that $c_p - c_v \sim 1$ cal/mole. °K combine to explain the enhanced value of the specific heat of nickel at high temperatures as compared with the value obtained from the law of Dulong and Petit.

The additional peak on the specific heat curve of nickel at 630°K is connected with the ferromagnetic transition which occurs at the Curie point. Similar peaks occur also when there is a change of atomic order in alloys, when the orientation of molecules in molecular crystals is destroyed, and so on. Peaks on the specific heat curve usually indicate the different orders of phase transition and are of great importance for revealing the existence of transitions in solids.

The Specific Heat and the Structure of Crystals

In crystals with pronounced anisotropy of structure and of binding forces, the vibrations of the atomic oscillators are anisotropic. Anisotropy of an oscillator leads to anisotropy in the vibration frequencies so only in a coordination crystal can the particles be thought of as isotropic oscillators with vibration frequencies which are the same in all three coordinate axes $\nu_1 = \nu_2 = \nu_3$. In a layer type crystal the vibration frequencies ν_1 and ν_2 parallel to the layer structure are the same and are much greater than the vibration frequency ν_3 perpendicular to the plane of the layer: $\nu_1 = \nu_2 \gg \nu_3$. In a chain type crystal the vibration frequency ν_1 parallel to the axis of the chain is greater than the vibration frequencies ν_2 and ν_3 in perpendicular directions; $\nu_1 \gg \nu_2 = \nu_3$.

In a molecular crystal it is necessary to distinguish the frequency of independent molecular vibrations ν_M and the frequency of intra-molecular vibrations ν_A. In the former case, the molecules vibrate as a whole, in the latter one can speak of the vibrations of individual atoms within the molecules. The frequencies of intra-molecular vibrations are very much greater than the frequencies of vibration of the molceules; $\nu_A \gg \nu_M$. The difference in character of the vibrations of the particles in structures of various types leads to peculiarities in the temperature dependence of the specific heat of crystals.

The Specific Heat of Layer-type and Chain-type Crystals

As Tarasov (1945) showed, the curve of the specific heat of layer- and chain-type crystals at low temperatures must differ from the T^3 law which applies to crystals with the coordination type of structure. In the limiting cases of layer- and chain-type crystals, when we can neglect interaction and flexibility of the layers and chains the specific heat at low temperatures must obey quadratic and linear laws: T^2 and T. Experimental data for the layer type crystals of arsenic and antimony at low temperatures follow a T^2 law well, while for the chain-type crystal of selenium there is a good approximation to the linear law.

The Specific Heat of Molecular Crystals

The specific heat of the molecular crystal of benzol has a complex temperature dependence (Fig. 232). The low temperature part of the curve shown in the figure corresponds to a low characteristic temperature and represents the excitation of molecular vibrations in the crystal. The abnormal increase in specific heat at temperatures above 80° K is connected with the excitation of intra-molecular vibrations. Saturation of these vibrations proceeds with a high characteristic temperature within the temperature range shown.

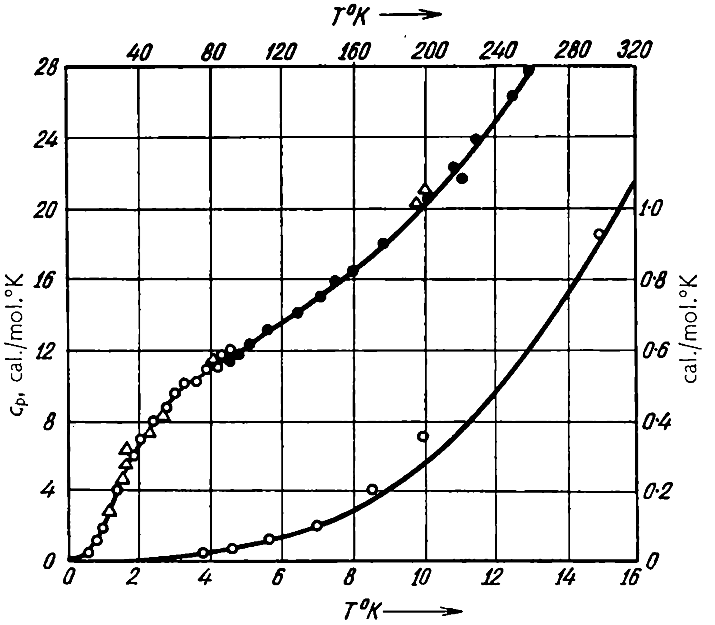


FIG. 232. Temperature dependence of the specific heat of benzol.

2. Principles of the Quantum Theory of the Specific Heat of Crystals

Energy and Specific Heat of a Quantum Oscillator

The sharp fall in specific heat of solids at low temperatures is a result of the breakdown of the classical law of equipartition of energy. The breakdown is itself a result of the quantum nature of the atomic oscillators (Einstein 1907). The energy spectrum of a quantum oscillator [*1.145*] $\varepsilon_n = nh\nu$ consists of a system of equally spaced levels (we need not take into account the zero point energy of an oscillator when calculating the specific heat). The vibration frequency of atoms in a solid $\approx 10^{13}$ sec^{-1} and corresponds to a frequency in the infra-red. The size of the step in the energy staircase of an atomic oscillator is $h\nu = 6 \times 10^{-14}$ ergs. At temperature $T = 300°$ K the mean thermal energy $kT = 4 \times 10^{-14}$ ergs which is comparable in order of magnitude with the energy of the quantum of vibration. Near to room temperature, almost all oscillators will be excited and will have thermal energy in accordance with the law of equipartition of energy. At low temperatures the mean thermal energy kT is so small ($kT < h\nu$) as to be insufficient to excite thermal vibrations of the oscillators. The majority of oscillators will thus remain in the state of zero point energy $E_0 = \frac{1}{2}h\nu$ and will execute only zero point oscillations, not thermal oscillations. Under these circumstances, only a small fraction of the atomic oscillators will be excited, and the mean energy of an oscillator will be less than kT. Thus at low temperatures the specific heat of a quantum oscillator will be less than the specific heat of a classical oscillator. To determine the temperature dependence of the specific heat we require the mean energy of a quantum oscillator.

The mean energy corresponding to thermal equilibrium in a quantum system with a discrete energy spectrum $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ can be determined as follows: the probability of finding a system in the nth state with energy ε_n is proportional to the Boltzmann factor $e^{-\varepsilon_n/kT}$. The probability of this state normalised to unity is

$$p_n = \frac{e^{-\frac{\varepsilon_n}{kT}}}{\sum\limits_n e^{-\frac{\varepsilon_n}{kT}}}. \qquad \ldots [\mathit{14.9}]$$

The quantity in the denominator of [*9*]

$$\sum_n e^{-\frac{\varepsilon_n}{kT}}, \qquad \ldots [\mathit{14.10}]$$

is the *Partition Function* of the system. The mean energy of the system is

$$\varepsilon_{\text{mean}} = \sum_n p_n \varepsilon_n = \frac{\sum\limits_n \varepsilon_n e^{-\frac{\varepsilon_n}{kT}}}{\sum\limits_n e^{-\frac{\varepsilon_n}{kT}}}. \qquad \ldots [\mathit{14.11}]$$

This expression can be put in the form

$$\varepsilon_{\text{mean}} = -\frac{d}{d\left(\frac{1}{kT}\right)} \ln\left(\sum_n e^{-\frac{\varepsilon_n}{kT}}\right), \qquad \ldots [\mathit{14.12}]$$

which shows that the mean energy of the system is the negative of the differential coefficient with respect to the reciprocal of temperature of the partition function. In a degenerate system more than one state may belong to a single energy level, in which case the partition function can be written in the form

$$\sum_n g_n e^{-\frac{\varepsilon_n}{kT}}, \qquad \ldots [\mathit{14.13}]$$

where g_n is the degeneracy of the level ε_n. The mean energy of the degenerate system is determined by the formula

$$\varepsilon_{\text{mean}} = -\frac{d}{d\left(\frac{1}{kT}\right)}\ln\left(\sum_n g_n e^{-\frac{\varepsilon_n}{kT}}\right), \qquad \ldots [14.14]$$

which is a generalisation of formula [*12*].

The partition function for an oscillator which is non-degenerate can be calculated from formula [*10*] by putting $\varepsilon_n = nh\nu$. Substituting

$$x = \frac{h\nu}{kT}, \qquad \ldots [14.15]$$

we have

$$\sum_n e^{-\frac{\varepsilon_n}{kT}} = \sum_n e^{-nx} = \frac{1}{1-e^{-x}}.$$

Putting this value for the partition function in [*12*], we obtain

$$\varepsilon_{\text{mean}} = -\frac{d}{d\left(\frac{1}{kT}\right)}\ln\frac{1}{1-e^{-x}} = -h\nu\frac{d}{dx}\ln\frac{1}{1-e^{-x}} = \frac{h\nu}{e^x-1}.$$

Replacing x by its value from [*15*] we obtain for the mean thermal energy of a quantum oscillator

$$\varepsilon_{\text{mean}} = \frac{h\nu}{e^{\frac{h\nu}{kT}}-1} = kT\,\frac{\frac{h\nu}{kT}}{e^{\frac{h\nu}{kT}}-1}. \qquad \ldots [14.16]$$

At high temperatures ($kT > h\nu$) $\varepsilon_{\text{mean}} \approx kT$ which is equal to the energy of a classical oscillator. At low temperatures we can neglect the one in the denominator of [*16*] and obtain

$$\varepsilon_{\text{mean}} = h\nu e^{-\frac{h\nu}{kT}} \quad (kT < h\nu). \qquad \ldots [14.17]$$

At low temperatures the energy of a quantum oscillator falls according to an exponential law and approaches zero asymptotically: $\varepsilon_{\text{mean}} \to 0$ as $T \to 0$.

Differentiating [*16*] with respect to temperature and multiplying by three we obtain the formula for the specific heat of a three dimensional quantum oscillator

$$c_a = \frac{d\varepsilon_{\text{mean}}}{dT} = 3k\left(\frac{h\nu}{kT}\right)^2 \frac{e^{\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}}-1\right)^2} = 3kf_E(x), \qquad \ldots [14.18]$$

where $f_E(x)$ is the Einstein specific heat function. At low temperatures

$$c_a = 3k\left(\frac{h\nu}{kT}\right)^2 e^{-\frac{h\nu}{kT}} \quad (kT < h\nu), \qquad \ldots [14.19]$$

and the temperature dependence of the specific heat is determined to a first approximation by the exponential. At $T = 0$; $c_a = 0$. The transition from the high temperature region in which the classical law is obeyed to the low temperature region where the quantum nature of the atomic vibrators makes itself felt is determined by equating the thermal energy kT to the energy of the quantum vibrator $h\nu$. Thus we obtain the Einstein characteristic temperature θ_E,

$$k\theta_E = h\nu. \qquad \ldots [14.20]$$

The characteristic frequency of vibration [*1.130*], $\nu = \frac{1}{2\pi}\sqrt{\frac{\alpha}{m}}$. Thus if the elastic coefficient α is constant, the characteristic temperature must fall as the mass of the oscillator increases; this agrees qualitatively with experiment (Table 66, series Be – Hg, Cu – Au). The exceptionally high value of the characteristic temperature of diamond ($\theta = 2000°$ K) which exceeds that of beryllium although the mass of the beryllium atom is less than the mass of the carbon atom can be explained by the high strength of binding in the diamond crystal (α is large).

Energy and Specific Heat of a System of Free Oscillators

Consider a system of N independent oscillators all vibrating with the same frequency ν. The vibrational energy of this system is

$$E = 3N\varepsilon_{\text{mean}}, \qquad \ldots [14.21]$$

and the vibrational molar heat is

$$c_v = 3Rf_E(x). \qquad \ldots [14.22]$$

The characteristic temperature is determined by formula [*20*]. Formulae [*20*]-[*22*] apply strictly speaking to a gas of vibrators. In a complete calculation of the energy of a gas we require to take account of the energy of translational motion of the oscillators. The application of these formulae to a crystal is possible only as a first approximation, so in spite of the achievement of the theory of specific heat of solids due to Einstein it does not agree quantitatively with experiment. At low temperatures, instead of the T^3 law, Einstein's theory gives an exponential dependence and a more rapid fall of specific heat than really takes place.

Energy and Specific Heat of a System of Coupled Oscillators

Molecules and crystals can be considered to be systems of coupled oscillators. If one of the oscillators is excited then because of the coupling, vibrations are started in neighbouring oscillators and vibration waves spread out in the system with a velocity which, in a crystal, is the velocity of sound. If the oscillators are excited thermally and the system is in thermal equilibrium, a system of stationary vibrations is set up which is a superposition of standing waves. The energy of elastic waves excited by thermal motion is determined by the amount of heat in the system and its temperature dependence—the specific heat of the crystal.

The most important difference between the system of coupled oscillators and a system of independent oscillators (a degenerate system) lies in the splitting up of the vibration frequencies leading to the removal of degeneracy. In place of a single frequency ν_0 which characterises the vibrations of independent oscillators we have now a spectrum of frequencies in a system of coupled oscillators. In a system of discrete particles this spectrum is made up of normal frequencies and in a homogeneous elastic body the spectrum is continuous.

1. Molecule. The vibration spectrum of molecules is discrete and consists of a comparatively small set of frequencies ν_i, the number of which is determined by the number of atoms in the molecule. The internal energy of a molecule is determined by the formula

$$E = \sum_{i=1}^{n} \varepsilon_{\text{mean}}(\nu_i) = \sum_{i=1}^{n} \frac{h\nu_i}{e^{\frac{h\nu_i}{kT}}-1}, \qquad \ldots [14.23]$$

which, when $\nu = \text{const.}$ and $\varepsilon_{\text{mean}} = \text{const.}$, becomes formula [*21*] for the degenerate system. Differentiating [*23*] with respect to T gives the vibrational part of the specific heat. For a gram-molecule we obtain

$$c_M = R \sum_{i=1}^{n} \left(\frac{h\nu_i}{kT}\right)^2 \frac{e^{\frac{h\nu_i}{kT}}}{\left(e^{\frac{h\nu_i}{kT}}-1\right)^2}. \qquad \ldots [14.24]$$

2. Crystal. Crystals are systems with a large number of degrees of freedom the vibration spectrum of which occupies a wide range of frequencies from $\nu_{\text{min}} \sim 10^4 \text{ sec}^{-1}$ to $\nu_{\text{max}} \sim 10^{13} \text{ sec}^{-1}$. The low frequency part of this spectrum is in the acoustic range and the high frequency part in the infra-red range. The crystal spectrum must be thought of as quasi-continuous, the distribution of frequencies within it being determined by the density of states function $q(\nu)$, where the number of vibrations in the frequency range from ν to $\nu + d\nu$ is $q(\nu)d\nu$. The density of states in the spectrum increases rapidly with increasing frequency. Therefore there is no great error if we take the spectrum to begin at $\nu_{\text{min}} = 0$. At the high frequency end of the spectrum the limiting frequency ν_m can be determined from the relation

$$\int_0^{\nu_m} q(\nu)\,d\nu = 3N. \qquad \ldots [14.25]$$

The internal energy of a crystal can be conveniently expressed in integral form

$$E = \int_0^{\nu_m} \varepsilon_{\text{mean}}\, q(\nu)\,d\nu = \int_0^{\nu_m} \frac{h\nu}{e^{\frac{h\nu}{kT}}-1}\, q(\nu)\,d\nu. \qquad \ldots [14.26]$$

The specific heat is obtained by differentiating [*25*] with respect to temperature

$$c_v = k \int_0^{\nu_m} \left(\frac{h\nu}{kT}\right)^2 \frac{e^{\frac{h\nu}{kT}}}{\left(e^{\frac{h\nu}{kT}}-1\right)^2}\, q(\nu)\,d\nu. \qquad \ldots [14.27]$$

The characteristic temperature θ_m is determined from formula [*20*] if in it we take the frequency ν to be the maximum frequency ν_m.

$$k\theta_m = h\nu_m. \qquad \ldots [14.28]$$

To find the energy and specific heat of a crystal from formulae [*26*] and [*27*], we must know the function $q(\nu)$. The basic problem of the theory of specific heats of solids is the determination of the frequency spectrum, that is, the determination of the form of the function $q(\nu)$, the density of elastic vibrational states and the limiting frequency ν_m.

Energy and Specific Heat of an Isotropic Solid

In the Debye theory of specific heat (1912), a crystal is considered as a continuous isotropic solid. The propagation of waves in a uniform medium is described by the wave equation

$$\nabla^2 u_i = \frac{1}{c_i^2}\frac{\partial^2 u_i}{\partial t^2} \qquad (i = l, t), \qquad \ldots [14.29]$$

where $u_i = u_i(x, y, z, t)$ is the displacement of points in the body, c_i the velocity of wave propagation. In a solid both longitudinal l and transverse t elastic waves can be propagated. The velocity of propagation of both waves is determined by the density ρ and the elastic properties of the medium

$$c_l = \sqrt{\frac{E}{\rho}}, \qquad c_t = \sqrt{\frac{G}{\rho}}, \qquad \ldots [14.30]$$

where E is Young's Modulus and G the shear modulus. As formula [*30*] shows, the velocity of propagation of waves in the continuous medium does not depend on the frequency and is the same for long and short waves. To determine the internal energy and specific heat of a solid we require to know the spectrum of these elastic vibrations. The simplest case is the vibration spectrum of a one-dimensional body, that is a thin rod.

1. The spectrum of characteristic vibrations of a thin rod. Consider an elastic rod of length L rigidly clamped at the ends at the points $x_1 = 0$ and $x_2 = L$. The solution of equation [*29*] for this case has the form

$$u_i(x, t) = A_n \sin 2\pi \frac{nx}{2L} \sin 2\pi\nu_n(i)\, t, \qquad \ldots [14.31]$$

$$(n = 1, 2, 3, \ldots;\ i = l, t)$$

where n is the number of the harmonic, A_n the amplitude and $\nu_n(i)$ the frequency. The characteristic vibrations of a rod form a system of standing waves which arise when there is a whole number n of half wavelengths in the length of the rod:

$$n\frac{\lambda_n}{2} = L.$$

This condition can be written in the form

$$\sigma_n = \frac{1}{\lambda_n} = \frac{n}{2L}, \qquad \ldots [14.32]$$

where σ_n is the wave number. The frequency of vibration is determined by the relation

$$\nu_n(i) = \frac{c_i}{\lambda_n} = \sigma_n c_i = \frac{c_i}{2L} n \qquad (i = l, t). \qquad \ldots [14.33]$$

The total number of vibrations Z_i in the interval from 0 to $\nu_n(i)$,

$$Z_i = \frac{2L}{c_i}\,\nu_n(i), \qquad \ldots [14.34]$$

is a linear function of frequency. Differentiating [*34*] we obtain

$$dZ_i = \frac{2L}{c_i}\,d\nu_n(i), \qquad \ldots [14.35]$$

from which

$$q_i(\nu) = \frac{2L}{c_i} = \text{const.} \qquad \ldots [14.36]$$

The density of distribution of vibrations in the spectrum of a one dimensional body is constant and does not depend on frequency.

2. *The spectrum of characteristic vibrations of a three-dimensional body.* The solution of a differential equation depends on the boundary conditions so the spectrum of characteristic vibrations of a solid should, strictly speaking, depend on the shape of the body. However the influence of the shape of the body affects only the low frequency parts of the vibration spectrum, that is those parts where the wavelength is comparable with the dimensions of the body. For wavelengths which are much less than the linear dimensions of the body, $\lambda_n \ll L$, the body can be considered as infinite and the distribution of energy in the short wavelength part of the spectrum as independent of the shape of the body. At normal temperatures the energy arising from the acoustic part of the spectrum must be very small since the internal energy and specific heat of a body are practically independent of shape. The dependence of the specific heat on the shape of the body can only manifest itself at very low temperatures when the relative weight of the low frequencies is large. It may also be evident in finely dispersed systems.

Leaving aside this limiting region of temperature we shall now calculate the spectrum of elastic vibrations of a body of the simplest and most symmetrical shape, that is to say a cube.

Consider a cube of edge A. Pick out one of the standing waves lying in a direction with direction cosines α, β, γ. A standing wave is formed on condition that along each cube edge we can measure out a whole number of projected half wavelengths

$$n_1\frac{\lambda_n}{2} = A\alpha; \quad n_2\frac{\lambda_n}{2} = A\beta; \quad n_3\frac{\lambda_n}{2} = A\gamma, \qquad \ldots [14.37]$$

where n_1, n_2, n_3 are positive whole numbers giving the order of the harmonic and $\lambda_n = \lambda\,(n_1, n_2, n_3)$ is the wavelength. From the relation connecting the direction cosines we obtain

$$(n_1^2+n_2^2+n_3^2)\,\frac{\lambda_n^2}{4} = A^2. \qquad \ldots [14.38]$$

Replace λ_n by the frequency of vibration according to formula [*33*] and we then obtain

$$n_1^2+n_2^2+n_3^2 = \frac{4A^2}{c_i^2}\,\nu_n^2(i) \qquad (i = l, t). \qquad \ldots [14.39]$$

Now construct a cubic space lattice with cube edge $\frac{c_i}{2A}$. The corners or nodes of this lattice, that is the points with integral coordinates, represent individual vibrations of the body. The nodes near to the origin represent the lower harmonics and those further away, the higher harmonics. The number of vibrations in the spectrum of a solid is limited by the minimum value of the wavelength, λ_{min} or the maximum value of the frequency, ν_{max}. For given values of $\nu_n(i)$ formula [*39*] is the equation of a sphere. All nodes inside this sphere represent possible vibrations which lie in the frequency interval from 0 to $\nu_n(i)$. The total number of vibrations of a solid, Z_i, is equal to the volume of one octant

$$Z_i = \frac{1}{8}\cdot\frac{4\pi}{3}\frac{8A^3}{c_i^3}\nu_n^3(i) = \frac{4\pi}{3}\frac{v}{c_i^3}\nu_n^3(i) \qquad (i = l, t), \qquad \ldots [14.40]$$

where $v = A^3$ is the volume of the body. Differentiating [*40*] we obtain

$$dZ_i = 4\pi\frac{v}{c_i^3}\nu_n^2(i) \qquad (i = l, t), \qquad \ldots [14.41]$$

whence

$$q[\nu_n(i)] = 4\pi\frac{v}{c_i^3}\nu_n^2(i) \qquad (i = l, t). \qquad \ldots [14.42]$$

Formulae [*40*] and [*42*] show that the number of vibrational modes of a three dimensional body is proportional to the cube of the frequency and the density of vibrational states to the square of the frequency. Similar calculations for a two dimensional body—a plate or membrane—leads to a square dependence for Z_i and a linear dependence for the density of vibrational states, as can easily be seen by comparison of the results for one-dimensional and three-dimensional cases.

The total number of vibrational states Z, taking account both of longitudinal and transverse waves will be

$$Z = Z_l + 2Z_t = \frac{4\pi v}{3}\left(\frac{1}{c_l^3}+\frac{2}{c_t^3}\right)\nu_m^3 = \frac{4\pi v}{c^3}\nu_m^3, \qquad \ldots [14.43]$$

where ν_m is the maximum frequency of vibration and c is the mean wave velocity determined by the relation

$$\frac{3}{c^3} = \frac{1}{c_l^3}+\frac{2}{c_t^3}. \qquad \ldots [14.44]$$

The complete distribution function, taking account both of longitudinal and transverse waves will be

$$q(\nu) = \frac{12\pi v}{c^3}\nu^2. \qquad \ldots [14.45]$$

Putting this value into [*26*] we obtain the formula for the internal energy

$$E = \frac{12\pi h v}{c^3}\int_0^{\nu_m}\frac{\nu^3\,d\nu}{e^{\frac{h\nu}{kT}}-1}. \qquad \ldots [14.46]$$

The vibration spectrum of a continuous solid is not limited at the high frequency end and must extend right to $\nu_m = \infty$, with the result that there is an infinite number of characteristic vibrations Z. In fact, the discrete atomic structure of solids limits the frequency of vibration to some value ν_m, which can be determined from [*43*] remembering that the maximum value $Z_{\max}$ is determined by the total number of degrees of freedom. For a system containing N particles the number of degrees of freedom is $Z_{\max} = 3N$. Using [*43*] and replacing Z by the number of particles, [*46*] becomes

$$E = \frac{9Nh}{\nu_m^3} \int_0^{\nu_m} \frac{\nu^3\, d\nu}{e^{\frac{h\nu}{kT}}-1}. \qquad \ldots [14.47]$$

Changing the variable of integration by means of [*15*] we obtain

$$E = 9Nh\nu_m \left(\frac{kT}{h\nu_m}\right)^4 \int_0^{x_m} \frac{x^3 dx}{e^x - 1}. \qquad \ldots [14.48]$$

Differentiating [*48*] gives the specific heat referred to the gram-molecule

$$c_v = 9nR\left(\frac{kT}{h\nu_m}\right)^3 \int_0^{x_m} \frac{e^x x^4 dx}{(e^x-1)^2} = 3nRf_D(x), \qquad \ldots [14.49]$$

where $f_D(x)$ is the Debye specific heat function, n the number of particles in the molecule. At low temperatures, $x_m \to \infty$, so in this region of temperature the upper limit of the integrals in [*48*] and [*49*] can be taken as ∞. In this case the value of the integrals does not depend on temperature. The internal energy of a solid becomes

$$E \approx (kT)^4 \qquad (kT < h\nu_m), \qquad \ldots [14.49a]$$

which is similar to the Stefan-Boltzmann law for radiation density. Thus we obtain the T^3 law for the specific heat at low temperatures. If we introduce the characteristic temperature from the Debye formula [*28*], then in the low temperature region, the internal energy and specific heat of a solid will be represented by the formulae

$$E = 3RT \times 19{\cdot}48\left(\frac{T}{\theta_m}\right)^3; \qquad c_v = 3R \times 77{\cdot}92\left(\frac{T}{\theta_m}\right); \qquad (kT < h\nu_m).$$

Fig. 233 shows curves of the specific heats $f_E(x)$ and $f_D(x)$, where the frequency of the Einstein oscillators ν has been taken equal to the Debye maximum frequency ν_m. The curve of f_E lies everywhere below the curve of f_D, as it must do since the Einstein theory takes no account of vibrations with frequencies $\nu < \nu_m$.

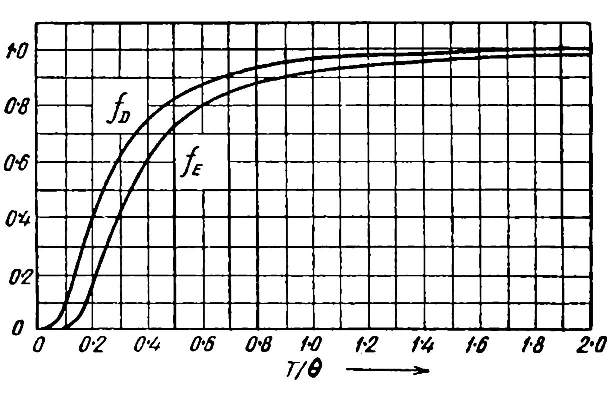


FIG. 233. Specific heat curves on the Einstein and Debye theories.

The Debye theory enables us to represent the characteristic temperature in terms of the elastic constant or the velocity of sound

$$\theta_m = \frac{h\nu_m}{k} = \frac{hc}{k\lambda_{\min}} = \frac{hc}{2ka}, \qquad \ldots [14.50]$$

where $\lambda_{\min} = 2a$ (a is the interatomic spacing). Taking the velocity of sound in a solid as $c = 3$ km/sec and $a \approx 3\times 10^{-8}$ cm, we obtain $\theta \approx 400°$ K, which agrees in order of magnitude with the experimental values.

If, in formula [*27*] we put the expressions for $q(\nu)$ applicable to infinitely thin plates and rods, we obtain results which are valid for molecular layer and chain type crystals respectively. The specific heats of these crystals depend respectively on T^2 and T at room temperatures. Tarasov generalised the Debye theory to the case of a uniform body in m dimensions. For the internal energy of a m-dimensional crystal he obtained

$$E(m) = 3mRT\left(\frac{T}{\theta}\right)^m \int_0^{x_{\max}} \frac{x^m\,dx}{e^x-1}, \qquad \ldots [14.51]$$

and for its specific heat

$$c_v(m) = 3m(m+1)R\left(\frac{T}{\theta}\right)^m \int_0^{x_{\max}} \frac{x^m\,dx}{e^x-1}\,3mR\left(\frac{\theta}{T}\right)e^{\frac{1}{\theta/T}}-1. \qquad \ldots [14.52]$$

The actual curves of the temperature variation of specific heat may have a very complex form depending on the structure of the crystal and the nature of the bonds between its structural elements. Thus, a layer type crystal, depending on the degree of binding between the layers, may have a continuous transition from a curve obeying the T^2 law to a curve obeying the T^3 law. Taking account of interaction between the layers, the specific heat of such a crystal can be written in the form

$$c(2-3) = D_2\left(\frac{\theta_2}{T}\right)-\left(\frac{\theta_3}{\theta_2}\right)^2\left[D_2\left(\frac{\theta_2}{T}\right)-D_3\left(\frac{\theta_3}{T}\right)\right]. \qquad \ldots [14.53]$$

If the binding between the layers is very small (the two-dimensional case), then the characteristic temperature θ_3 corresponding to interaction between the layers and connected with the exchange of elastic waves between them is zero; $\theta_3 = 0$ since $c_3 = 0$. In this case formula [*53*] reduces to its first term which represents the specific heat of a two-dimensional crystal. If the binding between the layers is as strong as that between atoms in each layer, then $\theta_3 = \theta_2$ and formula [*53*] becomes the same as that for a three-dimensional homogeneous crystal.

Another limiting case is that of molecular crystals, such as inorganic substances, with very weak interactions between the molecules. In the limit, such a crystal is of zero dimensions ($m = 0$), and the specific heat is given by the Einstein formula with constant frequencies which correspond to the frequencies of characteristic vibrations of the atoms in the molecule. The part of the specific heat which takes account of a weak interaction between the molecules can be determined by the Einstein formula or, for a homogeneous distribution of interaction, by the Debye formula. The formula for the specific heat of a molecular crystal containing N molecules and n atoms in each molecule can be expressed in the form

$$c_v = c_E + c_D, \qquad \ldots [14.54]$$

where

$$c_E = \sum_{i=1}^{3(n-1)} Rf_E\left(\frac{h\nu_i}{kT}\right) \quad \text{and} \quad c_D = \frac{R}{4n}\sum_{j=1}^{N}\int_0^{x_m} f_D\left(\frac{\theta_j}{T}\right) dx.$$

At low temperatures the second term is more important; at higher temperatures the first term, which corresponds to higher frequencies. This explains the bend which appears on the specific heat curves of molecular crystals (e.g. benzol).

Phonons

In the Debye theory, we have quantized the characteristic elastic vibrations of a solid which we considered as homogeneous. There was however nothing we could identify as vibrators. The energy of the elastic waves of the individual vibrations was taken as $\varepsilon = h\nu$, where h is the quantum constant and ν the frequency of vibration determined by formula [*33*]. In thermal equilibrium, the source of the energy of elastic waves in the crystal is thermal energy. The mean energy $\varepsilon_{\text{mean}}$ coming from a single vibration with frequency ν is determined by the Planck formula [*16*]. In a similar way as the quantization of electromagnetic radiation leads to the concept of photons, quantization of elastic waves leads to the concept of phonons. The sound quantum with frequency $\nu = c/\lambda$ (c is the velocity of sound) has energy $\varepsilon = h\nu$ and momentum $p = \varepsilon/c = h/\lambda$. From this point of view, the internal energy of a crystal is the energy of the phonons, the number of which increases with increasing temperature. At low temperatures the energy of the phonons is determined by the Debye formula [*49a*] which is analogous to the Stefan-Boltzmann law for the density of black body radiation. The difference between phonons and photons is that the T^4 law is applicable to photons at any temperatures while for phonons it is only applicable at temperatures less than θ_m. This difference is explained by the fact that as a result of the discrete structure of a solid, its spectrum is limited or cut off at the short wave end.

3. The Theory of Normal Vibrations of a Crystal

Born and Kármán (1912) solved the problem of the elastic vibrations of a crystal in the most strict and logical form taking account of the discrete periodic structure. The essential difference between the spectrum of normal vibrations of a crystal and the spectrum of vibrations of a homogeneous solid lies in the dispersion of the velocity of elastic waves in the discrete medium.

Characteristic Vibrations of a Monatomic Chain

Consider an infinite periodic chain of particles of mass m, arranged along the X-axis at equilibrium distances a from each other. Consider longitudinal vibrations of this chain. When the distances between the particles are changed, forces arise which tend to return the particles to their positions of equilibrium. For small displacements, these forces can be considered as quasi-elastic. Let ξ_l be the displacement of the lth particle along the X-axis. The equation of motion of any particle when displaced from its position of equilibrium can be written in the form

$$m\ddot{\xi}_l = -\alpha\,[(\xi_l-\xi_{l+1})-(\xi_{l-1}-\xi_l)], \qquad \ldots [14.55]$$

where α is the coefficient of quasi-elastic force. Rewrite equation [*55*] in the form

$$m\ddot{\xi}_l = -2\alpha\left(\xi_l-\frac{\xi_{l+1}+\xi_{l-1}}{2}\right), \qquad \ldots [14.56]$$

the solution of which has the form of a travelling sinusoidal wave

$$\xi_l = A_e^{2\pi i(\sigma x_l \pm \nu t)} \qquad (x_l = al;\ \ l = 1, 2, \ldots, N), \qquad \ldots [14.57]$$

the + and − signs of which correspond to two waves going in opposite directions. From [*57*] we have

$$\xi_{l+1}+\xi_{l-1} = \xi_l(e^{2\pi i\sigma a}+e^{-2\pi i\sigma a}) = 2\xi_l\cos 2\pi\sigma a. \qquad \ldots [14.58]$$

Differentiating [*57*] gives

$$\ddot{\xi}_l = -4\pi^2\nu^2\xi_l. \qquad \ldots [14.59]$$

Substituting [*58*] and [*59*] in [*56*] gives the following relation for the frequency of waves which can be propagated in the chain:

$$4\pi^2\nu^2 m = 2\alpha\,(1-\cos 2\pi\sigma a) = 4\alpha\sin^2\pi\sigma a,$$

whence

$$\nu = \frac{1}{\pi}\sqrt{\frac{\alpha}{m}}\sin\pi\sigma a = \nu_{\max}\sin\pi\sigma a. \qquad \ldots [14.60]$$

Formula [*60*] expresses the relation between the frequency of vibration ν, the wave number σ and quantities which determine the mechanical properties of the chain, that is, the lattice period a, the mass of the particle m and the elastic constant α.

The first essential difference between the discrete chain and the continuous rod is that the frequency is no longer directly proportional to the wave number σ. This is the phenomenon of dispersion: short waves to which correspond high frequencies of vibration are propagated more slowly than long waves because if the inertia of the particles. The velocity of propagation of waves in the chain is determined by the formula

$$v = \frac{\nu}{\sigma} = \nu_{\max}\frac{\sin\pi\sigma a}{\sigma}\cdot = a\sqrt{\frac{\alpha}{m}}\,\frac{\sin\pi\sigma a}{\pi\sigma a} = v_{\max}\frac{\sin\pi\sigma a}{\pi\sigma a}. \qquad \ldots [14.61]$$

The maximum velocity of propagation is for long waves. For $\sigma = 0$ we obtain

$$v_{\max} = a\sqrt{\frac{\alpha}{m}} = \sqrt{\frac{a\alpha}{\frac{m}{a}}} = \sqrt{\frac{\varepsilon}{\rho}}, \qquad \ldots [14.62]$$

where $a\alpha$ corresponds to Young's modulus E and m/a to the line density of the chain. Thus in the long wave limit, formula [*61*] is the same as formula [*30*] which expresses the velocity of propagation of waves in a continuous medium. The maximum frequency of vibration corresponds to the maximum value of $\sin \pi\sigma a = 1$, from which we obtain the maximum value of the wave number

$$\sigma_{\max} a = \tfrac{1}{2}; \qquad \sigma_{\max} = \frac{1}{2a}. \qquad \ldots [14.63]$$

Substituting this value for $\sigma_{\max}$ in [*61*] we obtain the velocity of propagation of the shortest waves in the chain

$$v_{\min} = \frac{2}{\pi} v_{\max} = 0{\cdot}64\, v_{\max}. \qquad \ldots [14.64]$$

The second essential difference between the chain and the rod lies in the cut-off of the frequency spectrum. As follows from [*60*], the maximum frequency of vibration of the particles of the chain

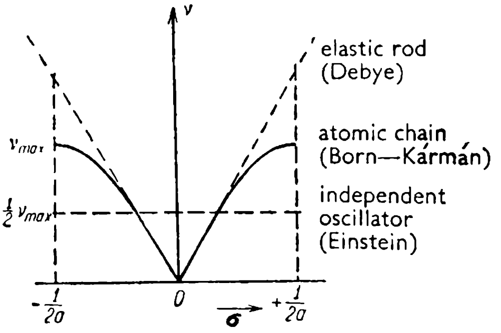
$$\nu_{\max} = \frac{1}{\pi}\sqrt{\frac{\alpha}{m}} \qquad \ldots [14.65]$$

is twice as large as the frequency ν_0 of an independent oscillator with mass m and elastic constant α. The maximum frequency of vibration corresponds to the maximum value of the wave number $\sigma_{\max}$ and the minimum value of the wave length which can be propagated in the chain is thus

$$\lambda_{\min} = 2a. \qquad \ldots [14.66]$$

This result has a simple interpretation. The shortest wavelength corresponds to vibrations of neighbouring particles with maximum phase difference which in this case is when neighbouring particles vibrate with opposite phase.

Fig. 234 shows the relation between the frequency ν and the wave number σ. For an atomic chain this is represented by two sinusoidal branches in the range of the variable from $-\frac{1}{2a}$ to $\frac{1}{2a}$. while for a rod or a string it is represented by a straight line which is tangential to the sine curve at the origin. This shows that in the long wave region the theory of the discrete structure of a solid agrees with the theory of the homogeneous medium. In the short wave region the two theories give different results. The horizontal dotted line in Fig. 234 represents the constant frequency of an independent oscillator as appears in the Einstein theory. We now proceed from the infinite chain to a finite chain containing N particles and having length $L = Na$. The limitation in length of the chain leads to a limitation of the vibration spectrum



ν
elastic rod (Debye)
atomic chain (Born—Kármán)
independent oscillator (Einstein)
ν_{max}
$\nu_0 = \frac{1}{2}\nu_{max}$
$-\frac{1}{2a}$
0
σ
$+\frac{1}{2a}$

FIG. 234. Dependence of frequency on wave number for a one-dimensional chain.

at the long wavelength end. If N is sufficiently large (this condition is satisfied even for microscopic crystals of length 1μ and containing several thousand atoms) the boundary conditions have little influence on the vibration spectrum. We can think of the finite chain as clamped at the ends corresponding to the boundary conditions of a clamped rod (see §2, page 389), or as a section of an infinite chain. In the latter case, the phases of vibration of the end atoms are taken to be the same. These are the periodic boundary conditions of Born and Kármán which, in an alternative form, consider the finite chain to have its ends connected together to form a circle. The finite length of the chain and its boundary conditions affect the result only at very low temperatures or in the case of very small crystals of colloid dimensions.

First consider the finite chain as being firmly clamped at the ends. The normal modes of the chain form a family of harmonics the wavelengths of which, as in the case of the rod, are determined by [*32*] which has the form

$$\lambda_n = \frac{2Na}{n}, \qquad (n = 1, 2, \ldots, N); \qquad \ldots [14.67]$$

the difference from the string is that the number of normal modes of the chain is limited and is equal to the number of degrees of freedom N. The shortest wavelength of the chain will be that of the highest harmonic

$$\lambda_{\min} = \lambda_N = 2a,$$

which agrees with [*66*]. The frequencies of vibration of a finite chain satisfy relation [*60*] in which we introduce the values from [*32*] and [*67*]. This gives

$$\nu_n = \nu_{\max} \sin \frac{\pi}{2} \frac{n}{N}. \qquad \ldots [14.68]$$

Differentiating [*68*] gives the frequency density

$$q(\nu) = \frac{dn}{d\nu} = \frac{2N}{\pi\sqrt{\nu_{\max}^2 - \nu^2}}. \qquad \ldots [14.69]$$

For an atomic chain, the density of vibrational states is not constant: the spectral density increases with increasing frequency of vibration and at

$$\sigma_{\max} = \frac{1}{2a} \quad q(\nu) \to \infty.$$

Characteristic Vibrations of a Diatomic Chain

Consider a periodic lattice consisting of particles with different masses m and M. Number all the particles from 1 to $2N$. The odd numbered particles are to have mass m and the even particles mass M. The distance between neighbouring particles is $a/2$ and the lattice period is a. The case we are considering corresponds to a one-dimensional crystal with a basis and the total length of the crystal $L = Na$ comprises N cells, each cell containing two particles. The equations of motion of the particles are of the form

$$\left.\begin{aligned} m\ddot{\xi}_{2l+1} &= -\alpha\left[(\xi_{2l+1}-\xi_{2l})-(\xi_{2l+2}-\xi_{2l+1})\right], \\ M\ddot{\xi}_{2l} &= -\alpha\left[(\xi_{2l}-\xi_{2l-1})-(\xi_{2l+1}-\xi_{2l})\right], \end{aligned}\right\} \qquad \ldots [14.70]$$

where ξ_{2l} and ξ_{2l+1} are the displacements of the particles from their equilibrium positions. These equations can be written in the form

$$\left.\begin{aligned} m\ddot{\xi}_{2l+1} &= -2\alpha\left(\xi_{2l+1}-\frac{\xi_{2l+2}+\xi_{2l}}{2}\right),\\ M\ddot{\xi}_{2l} &= -2\alpha\left(\xi_{2l}-\frac{\xi_{2l+1}+\xi_{2l-1}}{2}\right). \end{aligned}\right\} \qquad \ldots [14.71]$$

The solution of these equations is that of harmonic vibrations of particles in the form of running waves. As a result of difference of mass between the particles the amplitude of their vibrations will be different. Consider a wave travelling in the positive direction along the X-axis

$$\left.\begin{aligned} \xi_{2l+1} &= Ae^{2\pi i(\sigma x_{2l+1}-\nu t)}\\ \xi_{2l} &= Be^{2\pi i(\sigma x_{2l}-\nu t)} \end{aligned}\right\}. \qquad \ldots [14.72]$$

Here $x_{2l} = 2la$ and $x_{2l+1} = x_{2l}+\frac{1}{2}a$ are equilibrium position coordinates of particles and A and B are the respective amplitudes of vibration of light and heavy particles. The displacements of neighbouring particles are connected by the following relations

$$\xi_{2l} = \xi_{2l+1}\cdot\frac{B}{A}e^{-\pi i\sigma a};\quad \xi_{2l+2} = \xi_{2l+1}\cdot\frac{B}{A}e^{\pi i\sigma a};\quad \xi_{2l+1} = \xi_{2l}\cdot\frac{A}{B}e^{+\pi i\sigma a}. \quad \ldots [14.73]$$

Differentiating [72] twice gives

$$\ddot{\xi}_l = -4\pi^2\nu^2\xi_l. \qquad \ldots [14.74]$$

Substituting the value of the displacement ξ_l and the acceleration $\ddot{\xi}_l$ obtained from [72] in the equations of motion [71] we obtain the relation between the amplitudes of vibration of the particles:

$$\left.\begin{aligned} (2\pi^2 m\nu^2-\alpha)\,A+\alpha\cos\pi\sigma a\cdot B &= 0\\ \alpha\cos\pi\sigma a\cdot A+(2\pi^2 M\nu^2-\alpha)\cdot B &= 0 \end{aligned}\right\}. \qquad \ldots [14.75]$$

Consistency of equations [75] requires their determinant to vanish

$$\begin{vmatrix} 2\pi^2 m\nu^2-\alpha, & \alpha\cos\pi\sigma a\\ \alpha\cos\pi\sigma a, & 2\pi^2 M\nu^2-\alpha \end{vmatrix} = 0. \qquad \ldots [14.76]$$

Expanding this determinant we obtain the equation which gives the frequencies of vibration of the chain

$$4\pi^4\nu^4-\frac{\alpha(m+M)}{mM}2\pi^2\nu^2+\frac{\alpha^2}{mM}\sin^2\pi\sigma a = 0, \qquad \ldots [14.77]$$

the solution of which is

$$4\pi^2\nu^2 = \frac{\alpha(m+M)}{mM}\pm\frac{\alpha}{mM}\sqrt{(m+M)^2-4mM\sin^2\pi\sigma a}, \qquad \ldots [14.78]$$

which can be written in the form

$$\nu^2 = \frac{\alpha}{4\pi^2 mM}(m+M\pm\sqrt{m^2+M^2+2mM\cos 2\pi\sigma a}). \qquad \ldots [14.79]$$

Substituting the value of the vibration frequency in one or other of equations [75] we obtain the ratio of the amplitudes of vibration of the particles of different mass

$$\frac{B}{A} = -\frac{m-M\pm\sqrt{m^2+M^2+2mM\cos 2\pi\sigma a}}{2M\cos 2\pi\sigma a}. \qquad \ldots [14.80]$$

To a single value of the wave number there corresponds two different values of the frequency ν_- and ν_+ which arise from the different signs before the square root in [*79*]. Within the range of the variable σ from 0 to $\lambda/2a$, ν_- and ν_+ vary monotonically and take the following limiting values:

σ	ν_-	ν_+
0	0	$\nu_1+\nu_2$
$\frac{1}{2a}$	$\nu_1 = \frac{\alpha}{2\pi^2 M}$	$\nu_2 = \frac{\alpha}{2\pi^2 m}$.

The two solutions of Equation [*79*] form two frequency branches: the acoustical $\nu_- = \nu_-(\sigma)$ and the optical $\nu_+ = \nu_+(\sigma)$. These are shown in Fig. 235. When the wave number becomes small and the wavelength long, the frequency of vibration in the acoustical branch falls towards zero while in the optical branch it reaches a maximum. This difference is a result of the different ways in which neighbouring particles vibrate.

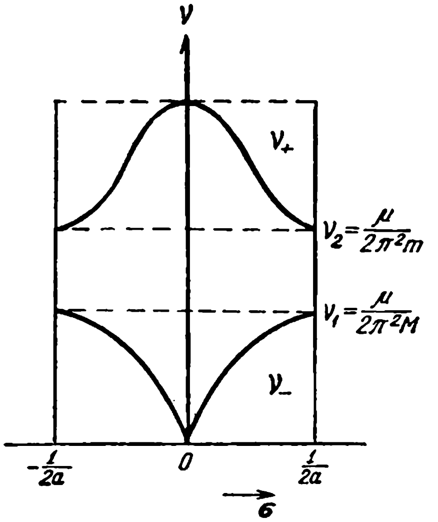


FIG. 235. Acoustical and optical branches of the vibration frequency spectrum of a diatomic chain.

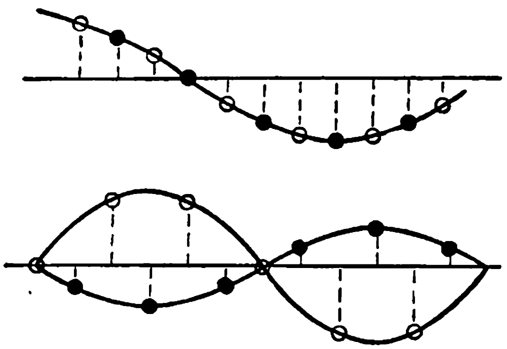


FIG. 236. Phases of vibration of the particles in the acoustic and optical wave.

In the acoustical branch, neighbouring particles vibrate with the same phase and long wavelengths correspond to movements of the unit cells as a whole (Fig. 236*a*); in the optical branch, neighbouring particles oscillate in antiphase and here long waves correspond to vibrations of the particles within the unit cells (Fig. 236*b*). These vibrations can be considered as vibrations of two translational lattices with respect to each other, each lattice consisting of particles of the same kind. When $\sigma = 0$ we obtain from [*80*]

$$\left(\frac{B}{A}\right)_- = 1; \qquad \left(\frac{B}{A}\right)_+ = -\frac{m}{M} \qquad (\sigma = 0), \qquad \ldots [14.81]$$

which confirms that the form of the vibrations of the chain is as described above. In the acoustical branch the particles vibrate with the same phases and the same amplitudes. In the optical branch the particles vibrate in antiphase. The amplitudes of vibration are reciprocally proportional to their masses. At the upper limit of the wave numbers $\sigma = \sigma_{\max}$ we have

$$\left(\frac{B}{A}\right)_- = 0; \qquad \left(\frac{B}{A}\right)_+ = -\frac{M-m}{m} \qquad (\sigma = \sigma_{\max}). \qquad \ldots [14.82]$$

In this case, in the acoustical branch, the heavier particles remain stationary.

The reality of the two branches in the vibration spectrum of crystals has recently been established by experiments on the scattering of cold neutrons.

The Vibrations of a Chain of Diatomic Molecules

Consider a chain consisting of atoms of all one kind but with different elastic constants α and β. Assume $\alpha > \beta$. The constant α is a measure of the force developed in intramolecular displacements while β is that of the force developed when the molecules move relative to each other. In this case the frequency of vibration is determined by the formula

$$\nu^2 = \frac{1}{4\pi^2 m}\left[2(\alpha+\beta) \pm \sqrt{4(\alpha+\beta)^2 - 16\alpha\beta \sin^2 \pi\sigma a}\right]. \qquad \ldots [14.83]$$

If $\alpha \gg \beta$, then the optical branch becomes a horizontal straight line with frequency

$$\nu_+ = \frac{1}{\pi}\sqrt{\frac{\alpha}{2m}}, \qquad \ldots [14.84]$$

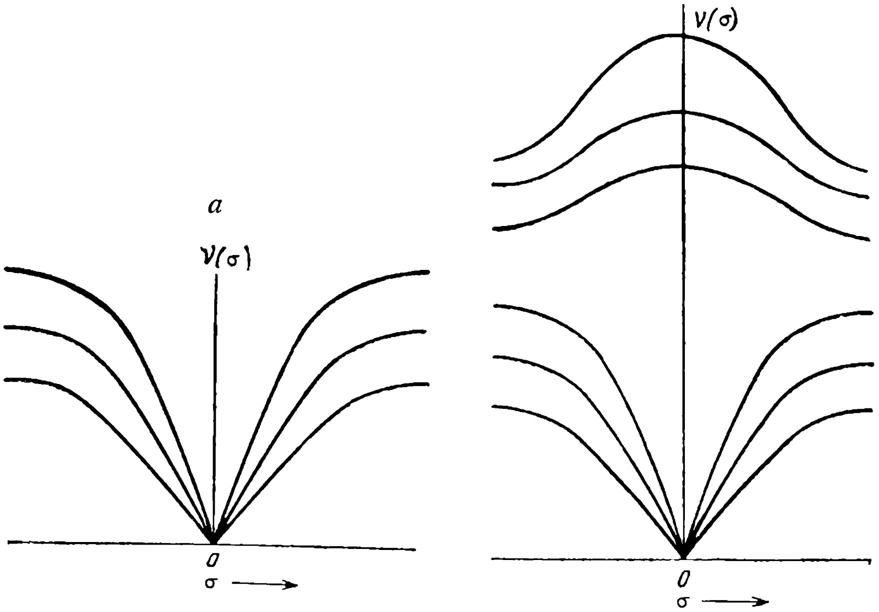


FIG. 237. Acoustic and optical branches of the spectrum of vibration of a three-dimensional crystal: (*a*) monatomic; (*b*) diatomic. This is a schematic representation; in fact the range of the variable σ is a three-dimensional zone. The curves shown represent the functions ν (σ) along straight lines through the origin.

which corresponds to the frequency of the characteristic vibrations of a diatomic molecule. The acoustical branch is determined by the formula

$$\nu_- = \frac{1}{\pi}\sqrt{\frac{\beta}{2m}}\sin \pi\sigma a, \qquad \ldots [14.85]$$

which corresponds to the vibrations of a chain consisting of particles with mass $2m$, that is the vibrations of a chain formed of rigid diatomic molecules. In the low frequency branch molecules vibrate as a whole, in the high frequency branch they vibrate as independent molecules. If the constants α and β are of similar magnitude then this division of the vibrations is not possible.

Vibrations of a Three-Dimensional Crystal

The theory of vibration of a one-dimensional chain can be generalised to the case of three dimensions. For a monatomic three-dimensional crystal we have three branches of the curves of $\nu(\sigma)$ as shown in Fig. 237*a*. For a diatomic crystal containing two kinds of atoms we have six branches (Fig. 237*b*) three of which correspond to acoustical vibrations and the three others to optical vibrations. Blackman obtained the distribution function for vibrational frequencies for a few simple lattices and hence determined the specific heat. In this work he took account of the interaction of atoms not only with nearest neighbours but also with atoms of the second coordination sphere. Fig. 238 shows curves of $q(\nu)$ on a relative scale for three branches $\nu(\sigma)$ (curves *b*, *c*, and *d*) and the composite curve *a* determined by an approximate method (the step curve). Calculations of the specific heat for halogen crystals obtained from Blackman's model do not agree very well with experiment.

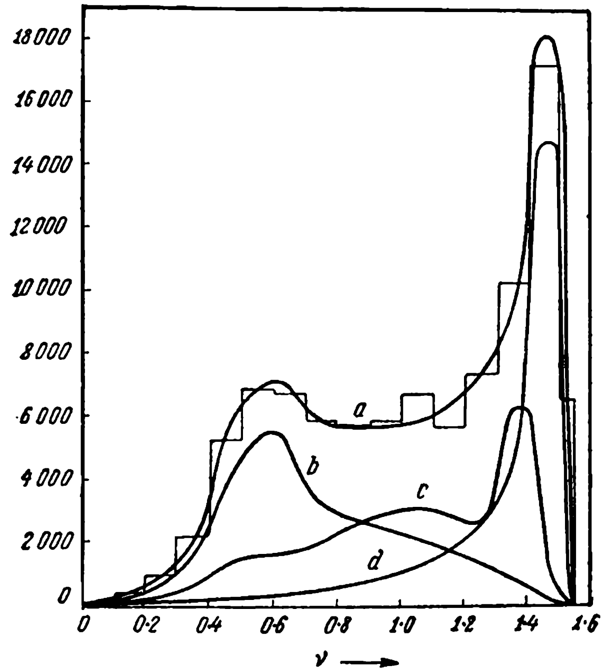


FIG. 238. Frequency distribution function for the vibrations of a simple cubic lattice.

4. Thermal Conductivity of Crystals

Definition

When a body is heated in an non-uniform manner the temperature tends to even out as a result of heat transfer from hotter to colder parts. Within the body there arise thermal currents the density of which $\frac{dQ}{dt}$ depends both on the character of the temperature field which is measured in an elementary region of the body by a temperature gradient $\frac{dT}{dx}$, and also on the properties of the medium determined by the coefficient of thermal conductivity κ

$$\frac{dQ}{dt} = \kappa \frac{dT}{dx} \Delta S. \qquad \ldots [14.86]$$

The thermal conductivity κ, which measures the ability of a medium to conduct heat is usually measured in the following units: cal/cm.degree.sec, erg/cm.degree.sec, W/cm.degree, the relation between which is determined by the following conversion:

$$1 \text{ Watt. sec} = 0{\cdot}239 \text{ cal} = 10^7 \text{ ergs.}$$

Phenomenologically thermal conductivity is described by a differential equation of the form [*1.13*] the solution of which depends on the initial and boundary conditions. We shall only be interested in one particular aspect of thermal conductivity, namely the mechanism of thermal conductivity in solids and the dependence of the coefficient of thermal conductivity on temperature, chemical nature and structure of the substance. Consider first some experimental facts relevant to the thermal conductivity of solids. We shall take account of the chemical nature of the substance by dividing all solids into metals and non-metals.

Thermal Conductivity of Non-metals

Thermal conductivity, in marked distinction to specific heat, is sensitive to small differences and defects in the atomic structure. This can be seen by the example of the temperature dependence of the specific heat and thermal conductivity of crystalline and amorphous substances. There is very little difference between the specific heats of a body in the crystalline and glassy states and the dependence on temperature is similar. It is very different with thermal conductivity: Fig. 239 shows curves of the temperature dependence of the thermal conductivity of some crystals and amorphous substances; curves *1* and *4* are for crystalline and fused quartz. As the temperature falls the thermal conductivity of quartz glass like its specific heat falls steadily while the thermal conductivity of crystalline quartz rises rapidly. At sufficiently low temperatures the thermal conductivities of crystalline and amorphous quartz differ by several orders of magnitude. The behaviour of other crystalline substances such as NaCl and CaF_2 (curves *2* and *3*) is similar to that of α-quartz, while typical glasses such as crown glass and Pyrex (curves *5* and *6*) behave similarly to quartz glass. The small value of the thermal conductivity at very low temperatures (point 7 which is for Jena glass) shows that as $T \to 0$ the thermal conductivity of amorphous substances, like the specific heat, tends to zero. To a first approximation the thermal conductivity

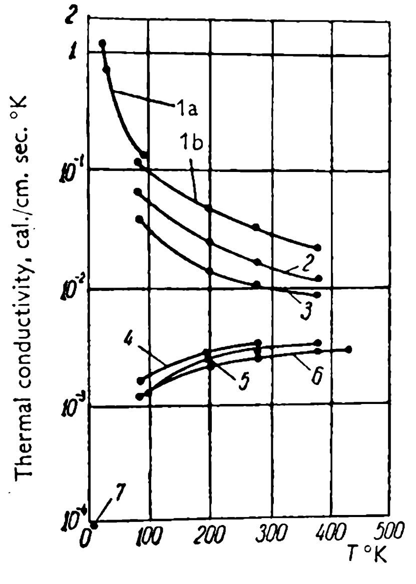


FIG. 239. Temperature dependence of the thermal conductivity of some crystals and amorphous substances.

of amorphous substances is proportional to the specific heat while the thermal conductivity of crystals at temperatures down to that of liquid air is reciprocally proportional to temperature. The behaviour of the thermal conductivity of crystals in the region of very low temperatures can be studied by investigations of the thermal conductivity of solids at the temperatures of liquid hydrogen and helium.

Fig. 240 shows the changes of the thermal conductivity of some crystals at low temperatures. For α-quartz the thermal conductivity rises as the specimen is cooled down to a temperature of the order of 10° K then the curve passes through a maximum and tends to zero as T tends to zero. For other crystals there is a similar behaviour: the curve $\kappa(T)$ passes through a maximum. The position, height and shape of the maximum depends on the composition and structure of the crystal. In the case of synthetic corundum α-Al_2O_3 (the crystals investigated were coloured, that is they were sapphires) the thermal conductivity maximum occurs at 40° K. Note specially the very high value of the thermal conductivity of corundum at its maximum. For a graphite crystal the thermal conductivity maximum occurs about 80° K. The position of this maximum is higher, the higher is the characteristic temperature of the crystal; being in the range $(\frac{1}{20}-\frac{1}{25})\,\theta$. If the thermal conductivity of graphite is measured in directions parallel to the atomic layers then the effective longitudinal vibrations are

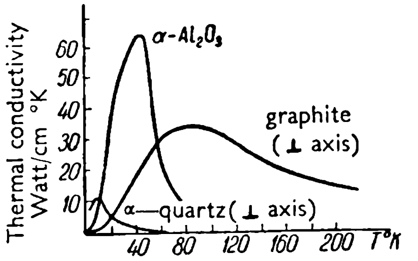


FIG. 240. The thermal conductivity of crystals at low temperatures.

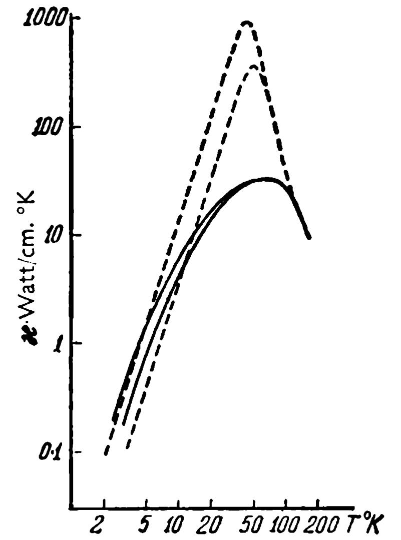


FIG. 241. Temperature dependence of the thermal conductivity of diamond.

those to which apply a characteristic temperature $\theta = 2100°$ K. This is almost equal to the characteristic temperature for diamond. Fig. 241 shows the curve $\kappa(T)$ for diamond, the maximum conductivity of which is 33 W/cm. °K occurring between 70 and 80° K. Because of the high characteristic temperature of diamond and graphite their thermal conductivities remain high above the temperature of the maximum and even at room temperature exceed those of silver and copper. Below the maximum, the thermal conductivity changes with temperature according to a T^3 law for graphite and to a T^n law for diamond where $n \sim 2{\cdot}7\text{-}2{\cdot}8$. These are approximately the same as the temperature dependence of the specific heat for layer and coordination crystals respectively. At very low temperatures the thermal conductivity of graphite, because of the anisotropy of its structure, falls rapidly and reaches values less than that for glasses. This is why graphite is a good thermal insulator at low temperatures.

The fact that the curve $\kappa(T)$ passes through a maximum suggests that there are different mechanisms of thermal conductivity of crystals in the temperature ranges above and below the maximum. The theory of the thermal conductivity of crystals

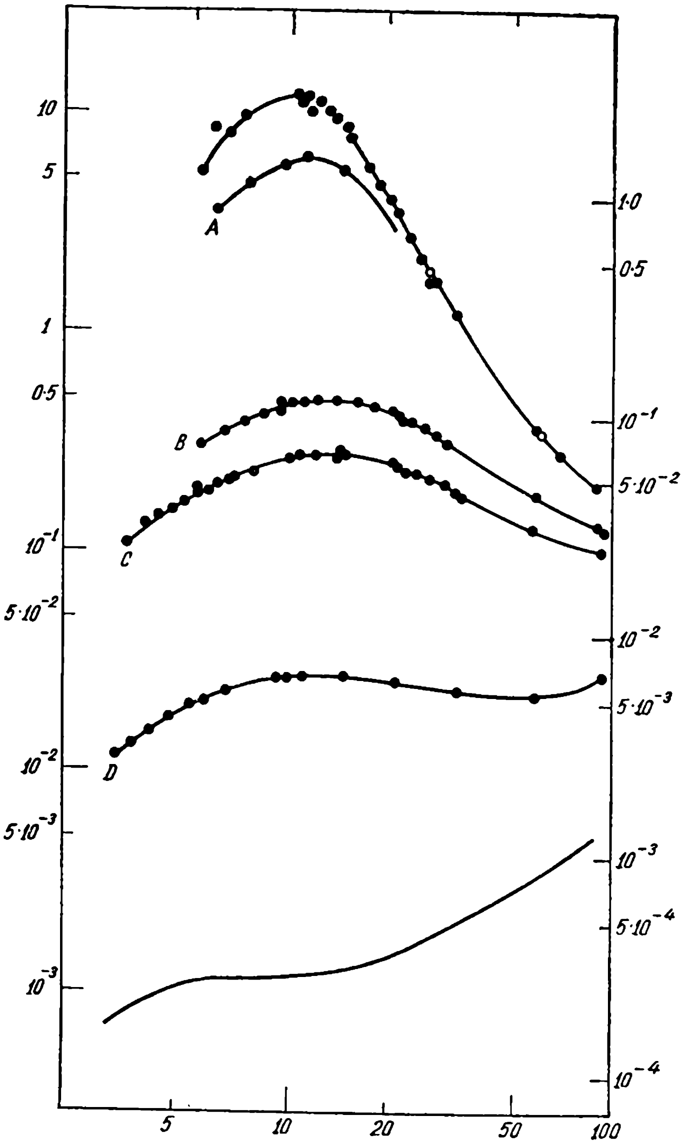


FIG. 242

Changes in the thermal conductivity of quartz, perpendicular to the c axis after neutron irradiation. The approximate doses corresponding to curves are A: 0·03, B: 1, C: 2·4, D: 19×10^{18} n/cm^2. The lowest curve gives values for an unirradiated specimen of quartz glass.

(Peierls, 1929) as will be shown below leads to the conclusion that at low temperatures the coefficient κ ceases to be a unique single-valued characteristic of a crystalline substance since the thermal conductivity begins to depend on the dimensions of the specimen. Experimentally, this dependence was demonstrated by de Haas and Biermasz (1935-38), in experiments on a crystal of potassium chloride. As the diameter of the specimen was reduced from 7·6 to 2·5 mm. κ fell according to an approximately linear law. Similar measurements with a diamond crystal were carried out by Berman, Simon and Ziman (1953). The diamond, of length 10·9 mm had a square cross section with side at first 3·9 mm which was progressively reduced to 3·2, 1·7 and finally 1·1 mm. The solid curves on Fig. 241 show the temperature dependence of the thermal conductivity of the specimens with largest and smallest cross sections. The thermal conductivity of the thinnest specimen is markedly less at low temperatures than that of the thickest specimen.

One can derive theoretical curves showing the high and low temperature branches of the $\kappa(T)$ curve of diamond and these are shown as dotted curves in Fig. 241. The theoretical value of the thermal conductivity of diamond at the maximum is about 1000 W/cm. °K, which is about 30 times higher than the measured value. This large discrepancy between theory and experiment can be explained by the presence of structural defects in the crystal which reduce its thermal conductivity. This conclusion is confirmed by measurements of the thermal conductivity of crystals which have been irradiated with neutrons. Fig. 242 shows the results obtained by Berman and others on quartz which has such a high sensitivity to neutron bombardment that after doses of $10^{22}n/\text{cm}^2$ its structure is completely amorphous. As the radiation doses increase, the thermal conductivity of α-quartz falls markedly and tends to the value for quartz glass.

Thermal Conductivity of Metals

At sufficiently high temperatures, the thermal conductivity of metals, like other crystalline substances, falls as the temperature rises (Fig. 243). There is a sharp drop in thermal conductivity at the transition from the crystalline to the liquid state, where long range order is lost and only short range order similar to that in amorphous

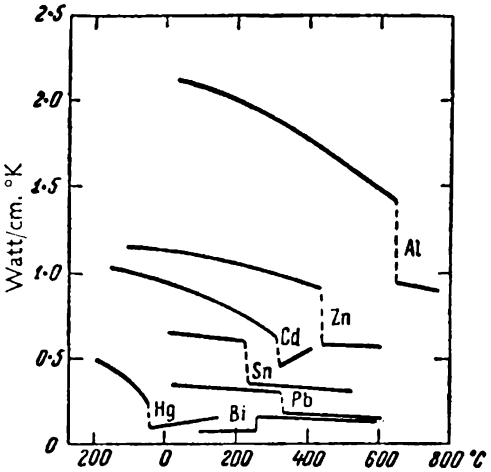


FIG. 243. Changes in the thermal conductivity of metals on melting.

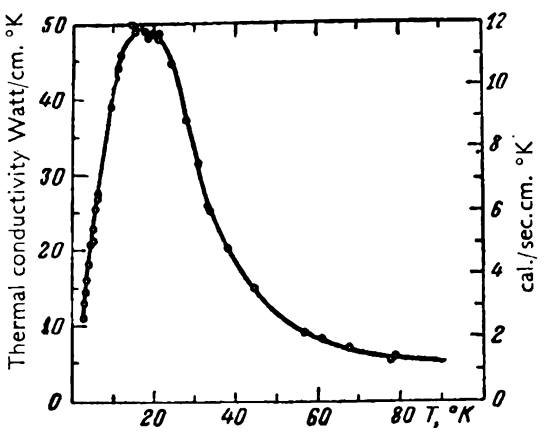


FIG. 244. Temperature dependence of thermal conductivity of copper.

substances is retained. Melting has thus a similar effect on the thermal conductivity of materials as other methods of making them amorphous but the change of thermal conductivity when metals melt is relatively less than when non-metals, for example quartz, are made amorphous. An exception is bismuth where on melting, the volume, thermal conductivity and electrical conductivity all increase. This can be explained by an increase in the metallic character of the bonds.

In the low temperature region, the thermal conductivity of metals, like that of non-metals, passes through a maximum and eventually tends to zero. This is illustrated by the results of Berman and MacDonald (1951) on the monovalent metals, copper (Fig. 244) and sodium. The sodium samples had different amounts of impurity; the purest had traces of silver while the less pure had traces of potassium and aluminium in amounts varying from 0·01 to 0·1%. The thermal conductivity of the less pure sodium reached a maximum of 14 W/cm. °K at 8° K. For the purest sodium, the thermal conductivity was still rising in that range of temperature up to values exceeding 30 W/cm. °K. Theoretical considerations show that maximum thermal conductivity for the purest sample should occur about 4·3° K, but this was beyond the experimental

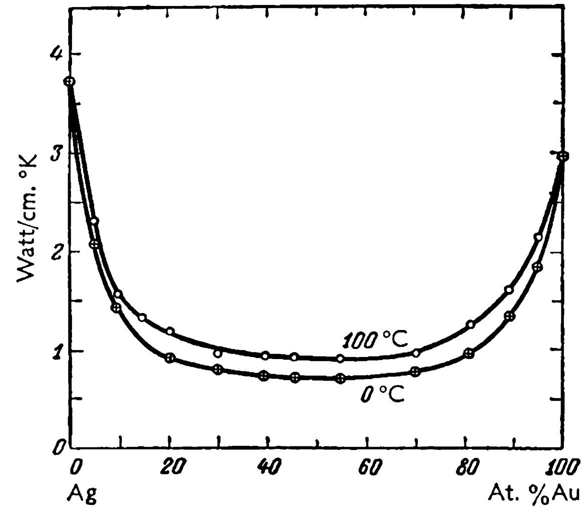


FIG. 245. Thermal conductivity of a solid solution.

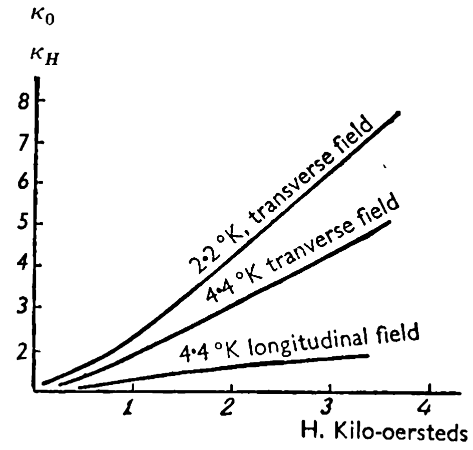


FIG. 246. Changes in the thermal conductivity of cadmium in a magnetic field.

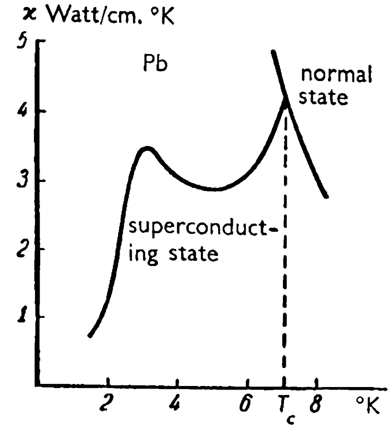


FIG. 247. Changes in the thermal conductivity of lead at the superconducting transition.

range. In solid solutions the thermal conductivity depends on concentration and falls as the amount of the second component increases. This can be seen by the example of the silver-gold alloys (Fig. 245).

In spite of the general similarity in the ways in which the thermal conductivities of metals and non-metals depend on temperature and the degree of perfection of the atomic structure, the mechanism of thermal conductivity is different for these two classes of substance. In metals, as distinct from non-metals, it is the conduction electrons which produce not only electrical conductivity but also most of the thermal conductivity. The part played by electrons in the thermal conductivity of metals can be directly demonstrated by the dependence of κ on magnetic field. Thus Fig. 246 shows that the thermal conductivity of cadmium falls by quite large factors in magnetic fields of the order of a few kilo-oersteds. As we would expect from consideration of the Lorentz force, a transverse field has a much greater effect than a longitudinal field lying in the direction in which the electrons move along the thermal current. As the temperature is lowered the influence of the magnetic field on the thermal conductivity increases. In superconductors, the transition to the superconducting state is accompanied by a sudden change in the slope of the curve and a fall in thermal conductivity since the superconducting electrons take no part in thermal exchange. Fig. 247 shows the curve $\kappa(T)$ for lead in the superconducting region.

The similarity between the mechanisms of physical and thermal conductivity of metals is shown by the Wiedemann-Frantz (1853) law which states that the ratio κ/σ is a constant for most metals at room temperature. Lorenz (1881) showed that the ratio

$$\frac{\kappa}{\sigma T} = L \qquad \ldots [14.87]$$

did not, to a first approximation, depend on temperature. The classical electron theory of metals gave for the Lorenz number L

$$L = 3\left(\frac{k}{e}\right)^2. \qquad \ldots [14.88]$$

The Wiedemann-Frantz law is satisfactorily established for metals at sufficiently high temperatures. At low temperatures it is found experimentally that L becomes smaller, and passes through a minimum. This can be seen for sodium in Fig. 248. The minimum is deeper when the impurity content is less. When there are no impurities, quantum theory shows that L should reduce to zero at absolute zero.

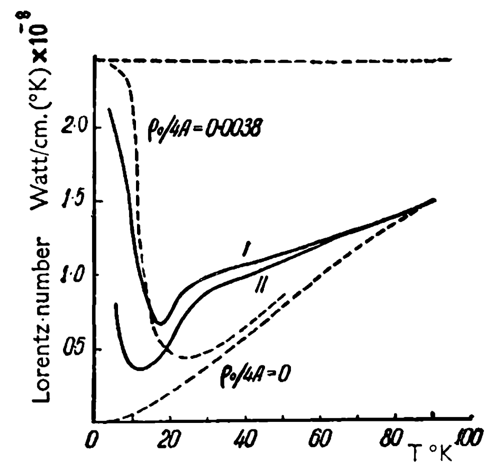


FIG. 248. The temperature dependence of the Lorenz number of sodium.

The Theory of Thermal Conductivity of Crystals

The thermal energy of a solid resides in the elastic vibrations of its particles. The velocity of propagation of these vibrations is, for long waves, equal to the velocity of sound, and is a few kilometres per second. As a result of dispersion, the velocity of propagation of short waves is reduced by about 40%, but nonetheless remains quite large in absolute magnitude. The time to propagate an elastic wave along a rod of length 25 cms is about 10^{-4} secs. On this reasoning the thermal conductivity of a solid should be large and the time of establishment of thermal equilibrium should be small. However experiment shows that this time is much larger than would follow from a simple calculation of the velocity of free propagation of elastic waves and the thermal conductivity is correspondingly less. As is well known, from the experience of glass blowing, a rod made of a non-metallic material can be heated in various parts and kept at high temperature for several hours without any marked rise in temperature in the unheated parts.

Debye (1914) explained this phenomenon by the scattering of thermal waves as they go through the crystal. The maximum energy in the thermal spectrum belongs to the short waves with wavelength comparable with the interatomic spacing. The shorter is the wavelength the greater is the scattering of waves on the inhomogeneities of the structure such as impurities, mosaic blocks and mechanical deformations (static stresses). Now wavelengths of sound waves are much greater than the dimensions of the atomic and microscopic inhomogeneities. Thus sound waves are propagated in solids without appreciable scattering, only suffering reflexion at the surfaces; for them, solids are transparent media. On the other hand, waves which correspond to the high frequency end of the spectrum are strongly scattered by microscopic or atomic inhomogeneities of the structure; the numbers of these inhomogeneities increasing in real crystals as their scale is reduced. For short waves crystals are like muddy materials in which the waves experience multiple diffusion scattering. This reduces their effective velocity of propagation in the same way as the scattering of particles lowers the velocity of diffusion in gases, although the absolute magnitude of their velocity of free motion is still large.

The scattering of thermal waves on static stresses in the structure, or their reflexion on the surface of crystals is insufficient to explain thermal conductivity. It can be concluded from analysis of the dependence of thermal conductivity on the concentration and on the degree of plastic deformation that a large part of the scattering of thermal waves must still take place in an ideal crystal without static stresses. This scattering can take place only on crystal inhomogeneities produced by thermal motion (dynamic stresses). Consequently thermal waves must be able to scatter each other. But if we think of these waves as produced by the movements of particles which are subject to Hooke's law then they ought to obey the principle of superposition: the waves ought to proceed independent of each other.

To explain the possibility of redistribution of energy between the various frequencies and the establishment of thermal equilibrium in a crystal it is necessary to assume non-linear thermal vibrations of the atoms, that is the breakdown of the law of proportionality between the displacement of the particles from their positions of equilibrium ξ produced by thermal motion and the restoring force $f(\xi)$

$$f(\xi) = \beta\xi - g\xi^2 + h\xi^3 - \dots \qquad \dots [14.89]$$

The expression for the potential energy of an atomic vibrator becomes

$$U(\xi) = \tfrac{1}{2}\beta\xi^2 - \tfrac{1}{3}g\xi^3 + \tfrac{1}{4}h\xi^4 - \dots, \qquad \dots [14.90]$$

which is no longer simply quadratic. Thus we have to take account of anharmonicity of thermal vibrations in a solid.

The anharmonic terms—those with coefficients in $g, h, \dots$ are larger, the larger are the displacements of the atoms from their positions of equilibrium. The amplitudes of thermal vibration in solids ξ are much larger than the displacements of atoms produced by macroscopic elastic deformations which obey Hooke's law. Such deformations do not exceed 0·1% which corresponds to $\xi_{\mathrm{el}} = 0{\cdot}003$ Å. The amplitudes of thermal vibration in crystals at room temperature as can be seen from X-ray data are $\xi_{\mathrm{th}} \simeq 0{\cdot}10$-$0{\cdot}15$ Å, which is about 3-5% of the interatomic spacing. As the temperature is lowered the amplitudes of thermal vibrations become less. However, even at absolute zero the zero point vibrations remain in the crystal. The energy of zero point oscillations $\frac{1}{2}h\nu$ corresponds to the energy of thermal vibration at a temperature $T = \frac{1}{2}\theta$, where θ is the characteristic temperature. This shows that the zero point oscillations will also be anharmonic. Thermal vibrations including the zero point oscillations are anharmonic in all regions of temperature in which the crystal exists.

Debye put forward the theory of thermal conductivity of a solid considering it as a homogeneous elastic medium, in which the thermal resistivity is due to scattering of elastic waves in a similar manner to the scattering of light in translucent media which is described by Bouguer's or Lambert's law. Thus when the waves pass through a length x in the medium their intensity falls according to an exponential law $I = I_0 e^{-\mu x}$, where μ is the scattering coefficient and depends on the wavelength and on the scattering density (non-transparency) of the medium. Introducing a parameter which measures the scattering of the waves, that is the path length $l = 1/\mu$ in which the intensity of the wave falls by a factor e, Debye obtained a formula for the thermal conductivity of a solid,

$$\kappa = \frac{cv}{4\mu} = \tfrac{1}{4}cvl, \qquad \ldots [14.91]$$

where c is the volume specific heat. Calculation of the scattering coefficient showed that at sufficiently high temperatures μ was proportional to g^2 where g is one of our anharmonic coefficients. The velocity of sound does not depend much on temperature so, in the region of temperature where the specific heat is constant, the thermal conductivity is reciprocally proportional to temperature. This agrees qualitatively with experiments on crystals at temperatures above the thermal conductivity maximum.

Formula [*91*] is similar to the formula for the thermal conductivity of gases

$$\kappa = \tfrac{1}{3}cvl \qquad \ldots [14.92]$$

in the kinetic theory of gases. Formula [*92*] can be applied to the solid if we take the thermal energy and thermal conductivity to be the energy and thermal conductivity of the phonon gas. The quantity l in this case is the mean free path of a phonon. The different numerical coefficients in [*91*] and [*92*] are not important since the theory of thermal conductivity of crystals does not as yet enable us to calculate the absolute value of κ. Phonon scattering satisfies the basic laws of conservation of energy and momentum

$$h\nu_1 = h\nu_2 + h\nu_3, \qquad \ldots [14.93]$$

$$\mathbf{p}_1 = \mathbf{p}_2 + \mathbf{p}_3, \qquad \ldots [14.94]$$

and leads to the establishment of thermal equilibrium. From [*94*] it follows that the direction of the thermal current is not lost on scattering and consequently the thermal resistivity of a body should be zero. Thermal resistivity arises as a result of scattering of phonons by processes which can conditionally be called collisions of phonons with the crystal lattice. As Peierls showed, equation [*94*] in this case has a different form

$$\mathbf{p}_1 = \mathbf{p}_2 + \mathbf{p}_3 + \frac{h}{2\pi a}\mathbf{q}, \qquad \ldots [14.95]$$

where a is the lattice spacing, and q is the unit vector the possible directions of which are determined by the lattice symmetry. The last term on the right hand of [*95*] determines the quantity of momentum $\frac{h}{2\pi a}$, transferred as a result of lattice scattering. Such collisions (scattering processes) were called by Peierls 'Umklapp' or 'turn-round' processes. They are responsible for the thermal resistivity of ideal crystals. The thermal conductivity of a crystal arising from the motions of phonons with given frequency is described by formula [*92*], where l is the mean free path of a phonon between two Umklapp processes. Taking into account the frequency spectrum of crystals and polarisation of the phonons corresponding to transverse and longitudinal vibrations, formula [*92*] becomes

$$\kappa = \tfrac{1}{3} \sum_{k_i} \sum_{j} c(k_{i,j}) v(k_{i,j}) l(k_{i,j}), \qquad \ldots [14.96]$$

$$(j = t, l),$$

where j determines the polarisation of the phonon and k_i its frequency. At high temperatures the contribution to the thermal conductivity $c(k_i, j)$ of phonons of different frequencies is the same. The specific heat is also constant. The mean free path l of a phonon depends on its excess energy and on the number of phonons which at high temperatures is proportional to T. Thus l and κ are reciprocally proportional to temperature as we have already shown for the elastic wave model.

From [*95*] it follows that phonons can take part in Umklapp processes if their momentum is not less than $\frac{h}{2\pi a}$, and this gives the lower limit of the energy of these phonons:

$$h\nu' = \frac{hv}{2\pi a}. \qquad \ldots [14.97]$$

Since

$$h\nu_{\max} = \frac{hv}{\lambda_{\min}} = \frac{hv}{2a},$$

we have

$$h\nu' \simeq \tfrac{1}{3} h\nu_{\max} = \tfrac{1}{3} k\theta. \qquad \ldots [14.98]$$

On Peierls' calculations the numerical factor in [*98*] is $\frac{1}{2}$. The mean energy of these phonons is

$$\varepsilon' = \frac{\frac{1}{3} k\theta}{e^{\theta/3T} - 1}.$$

At low temperatures this energy is proportional to $e^{-\theta/3T}$. Thus thermal resistivity should be proportional to $f(T)e^{-\theta/3T}$ where $f(T)$ is the temperature dependence of the specific heat. These considerations show that the increase of thermal conductivity of a crystal according to a hyperbolic law changes to a more complex law in which the most important factor is the exponential.

At sufficiently low temperatures, the density of the phonon gas falls, thus the major factor begins to be the collisions of phonons with the walls, that is with the crystal boundaries. When these collisions are predominant, the thermal conductivity begins to depend on the shape of the specimen. This corresponds to the transition from a dense phonon gas to a rarified gas and to the vacuum state. Neglecting completely the collisions of phonons with the lattice and considering only collisions with the walls, Casimir found that the thermal current in a cylindrical specimen of circular or square cross section was proportional to the temperature gradient, to the cube of the absolute temperature and the cube of the radius (or the cube of the side). If we determine from this the thermal conductivity κ', then it will be proportional to the radius of the crystal and the cube of the temperature

$$\kappa' \sim rT^3. \qquad \ldots [14.99]$$

Formula [*99*] explains qualitatively the fall of thermal conductivity at $T \to 0$ and its dependence on the dimensions of the specimen.

Debye explained the great difference in thermal conductivity between crystals and amorphous substances by the differences in mean free path of the thermal waves. The experimental value of the thermal conductivity enables us to obtain the mean free path: For quartz glass at room temperature $l \sim 8$ Å, which is comparable with the dimensions of the unit cell of α-quartz. Thus phonons are scattered on static lattice defects of atomic scale. The mean free path of the phonons in amorphous substances does not depend on temperature so in this case the thermal conductivity must be proportional to the specific heat.

The mechanism of thermal conductivity in metals is more complex than in non-metals since heat is carried by conduction electrons as well as by thermal waves. The total thermal conductivity of a metal is made up of two parts; the thermal conductivity of the phonon gas κ_{ph} and the thermal conductivity of the electron gas κ_e

$$\kappa = \kappa_{ph} + \kappa_e. \qquad \ldots [14.100]$$

Comparison of the thermal conductivities of metals and non-metals shows that $\kappa_{ph} \sim \kappa_e$. This being so, it still remains unclear why the majority of metals obey the Wiedeman-Frantz law with fair accuracy over a wide range of temperature with the exception of the low temperature region. The explanation for this might be that the presence of free electrons in a metal is a source of additional strong scattering of phonons and reduces the importance of scattering of phonons by thermal inhomogeneities of the lattice. Thus in pure metals the thermal conductivity of the lattice κ_{ph} is much less than κ_e, being in fact about 1-2% of the electron conductivity. In alloys, the role of the thermal conductivity of the lattice may become more important because of additional scattering of phonons on the impurities.

5. Thermal Expansion of Crystals

Anharmonic Thermal Vibrations

The dimensional changes of solids and liquids on heating are connected with differences in the behaviour of the attractive and repulsive forces when the atoms are moved from their positions of equilibrium. The potential energy curve for the interaction of two particles (Fig. 47) has a clearly asymmetrical shape relative to the vertical ordinate through the minimum of the curve. When atoms are brought closer together, the repulsive energy increases more rapidly than does the attractive energy when the atoms are taken further apart. When the amplitude of thermal vibrations is increased the minimum spacing between the atoms is reduced and the maximum spacing increased so, by the above reasoning, the increase of spacing when the atoms separate is greater than the reduction of spacing when they come together. This is why the dimensions of a body increase on heating. In an anisotropic crystal the asymmetry of the curve of interaction potential may be the other way round for some directions. There may then be in these directions a contraction of linear dimensions on heating. If the potential energy curve were symmetrical, for example a parabola (as in harmonic oscillators), there would be no change in dimensions of the body when the temperature is changed. The lowest part of the potential curve close to the minimum, can always be considered to be a parabola with sufficient accuracy. This explains why the thermal expansion coefficient of all bodies near absolute zero tends to zero. Thus the reason for thermal expansion is found in the anharmonicity of thermal vibrations of atoms in the crystal. This is also applicable to molecules where the mean interatomic spacing likewise depends on temperature.

The Thermal Expansion Tensor

Consider the change in dimensions of a crystal on heating (thermal deformation) at constant pressure. The thermal expansion of the majority of crystals is anisotropic and must be represented by a polar diagram of the linear expansion coefficient. The coefficient of linear thermal expansion

$$\alpha = \frac{1}{l}\frac{dl}{dT}, \qquad \ldots [14.101]$$

where l is the length of the crystal at temperature T, dl the thermal deformation produced by a change of temperature dT. The change of dimensions of crystals on heating from temperature T_1 to T_2 is given by an exponential formula

$$l = l_0 e^{\int_{T_1}^{T_2} \alpha(T)\,dT}, \qquad \ldots [14.102]$$

which when α is a constant takes the simpler form

$$l = l_0 e^{\alpha(T_2 - T_1)}. \qquad \ldots [14.103]$$

Since the coefficients of linear thermal expansion of solids are small, of order 10^{-5}/°K, [*103*] reduces to a linear law of expansion $l \simeq l_0\,[1+\alpha(T_2-T_1)]$, which is valid over small ranges of temperature.

When a crystal is uniformly heated, all its parts expand uniformly and as a result the crystal experiences a *homogeneous* deformation produced by a change of the scalar quantity T. Because of anisotropy, the coefficients of linear expansion of crystals of all systems except the cubic depend on direction. Thermal deformation is described by the tensor of thermal expansion. Take α_{11}, α_{22}, α_{33} as the principal linear coefficients along three mutually perpendicular axes. If to start with the crystal has a spherical shape with unit radius then after heating by 1° C the crystal will have an ellipsoidal shape,

$$\frac{x_1^2}{(1+\alpha_{11})^2} + \frac{x_2^2}{(1+\alpha_{22})^2} + \frac{x_3^2}{(1+\alpha_{33})^2} = 1, \qquad \ldots [14.104]$$

which represents the tensor of thermal expansion. Since the coefficients of expansion are small, the equation of the thermal expansion ellipsoid can be put in the form

$$x_1^2+x_2^2+x_3^2 = 1+2(\alpha_{11}x_1^2+\alpha_{22}x_2^2+\alpha_{33}x_3^2). \qquad \dots [14.105]$$

The radius vector of the ellipsoid can be written in the form $r = 1+g$, where g is the coefficient of linear expansion of the crystal in the direction of the radius vector. It is clear that

$$r^2 = x_1^2+x_2^2+x_3^2 = (1+g)^2 \approx 1+2g. \qquad \dots [14.106]$$

Comparing [*105*] and [*106*] we obtain

$$g = \alpha_{11}x_1^2+\alpha_{22}x_2^2+\alpha_{33}x_3^2. \qquad \dots [14.107]$$

Since the radius vectors of the ellipsoid of thermal expansion are close to unity, the coordinates x_1, x_2, x_3 are equal to the cosines of the angles between the radius vector and the coordinate axis. Supposing that the radius vector coincides with the direction of a new axis which we call X_3 the equation for the direction cosines for it can be written in the form $x_1 = c_{13}$; $x_2 = c_{23}$; $x_3 = c_{33}$. Finally the coefficient of linear expansion of a crystal in an arbitrary direction appears as

$$g = \alpha_{11}c_{13}^2+\alpha_{22}c_{23}^2+\alpha_{33}c_{33}^2. \qquad \dots [14.108]$$

For a cubic crystal $\alpha_{11} = \alpha_{22} = \alpha_{33}$, $g = \text{const}$, and the ellipsoid of thermal expansion, as for an isotropic body, reduces to a sphere. For crystals of other systems which have an axis of symmetry with multiplicity greater than C_1, the ellipsoid of thermal expansion is oriented relative to the axis of symmetry as shown above.

In uniaxial crystals (tetragonal, hexagonal and rhombohedral) the ellipsoid of thermal expansion is an ellipsoid of revolution with axis of revolution parallel to the principal axis of the crystal. In crystals of other systems the ellipsoid of revolution is triaxial. In rhombic crystals the principal axes of the ellipsoid of thermal expansion are parallel to the diad axes and its position is simply determined relative to the crystallographic axes. In monoclinic crystals one of the principal axes of the tensor of thermal expansion is parallel to the monoclinic axis, and the two others lie in the monoclinic plane, their position relative to the crystal axis being determined by means of the monoclinic angle. In triclinic crystals the position of the ellipsoid of thermal expansion cannot be immediately connected with the crystallographic axis and is determined by means of the three angles.

Experimental Results

Experimental values of the coefficients of thermal expansion for crystals of various systems are given in Table 67.

The smallest coefficient of linear expansion is that for diamond, the carbon atoms of which are linked together by strong covalent bonds. The largest value is that for the organic molecular crystal of para-nitroaniline. In uranium and calcite one of the coefficients of linear expansion is negative.

Table 68 gives values of the coefficients of linear expansion for various cubic metals at a variety of temperatures.

As the temperature is lowered the coefficients of linear expansion become less and as the temperature approaches absolute zero the coefficients tend to zero. Thermal expansion depends on the shape of the potential energy curve and on the level of

Table 67

Coefficients of Thermal Expansion of Some Crystals at Room Temperature (10^6 degree^{-1})

Crystal symmetry	Crystal	α_{11}	α_{22}	α_{33}
Cubic	Diamond	0·6		
Tetragonal	β-Sn	30·5		15·5
Hexagonal	α-Quartz	13·7		7·5
Rhombohedral	Calcite	−5·56		24·9
Rhombic	α-U	21·7	−1·5	23·2
Monoclinic	para-nitroaniline	150·3	7·5	24·0

Table 68

Temperature Dependence of the Coefficients of Linear Expansion $\alpha\,.\,10^6$ degree^{-1} for Various Cubic Metals

Metal	Coefficients of linear expansion at the following temperatures.							
	−200°	−100°	0°	100°	200°	400°	600°	800°
Al	7	19	22·5	24·5	26·5	30	31	—
Fe	3	10	11·6	12·7	13·5	14·8	16·0	—
W	—	4·2	4·3	4·5	4·7	4·9	5·2	5·5

thermal energy. Near to absolute zero, thermal excitation of the quantum vibrators of a solid is very small which explains the trend $\alpha \to 0$ as $T \to 0$. The temperature dependence of the coefficient of linear expansion for metals is thus similar to the temperature dependence of the specific heat. The connection between α and c_v is given by the thermodynamic relation

$$\alpha = \frac{\gamma \kappa c_v}{3v}, \qquad \ldots [14.109]$$

which is known as Grüneisen's law; here κ is the coefficient of compressibility, and γ a constant the value of which varies between 1·5-2·5 for different metals.

Thermal Expansion and Crystal Structure

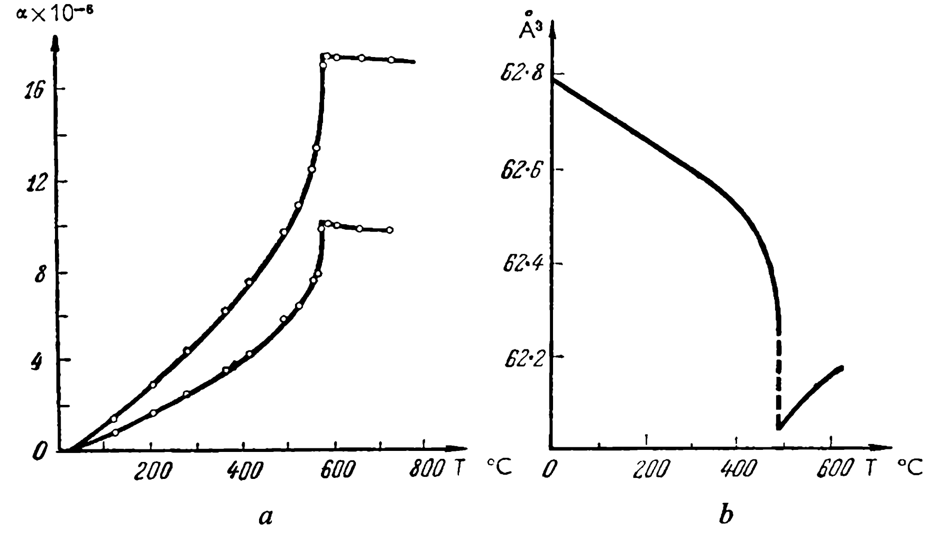
The structural properties of a crystal are reflected by regularities in its thermal expansion. This can be seen by the example of crystals with layer and chain-like structures. As an example of a layer-type crystal consider hexagonal graphite and of a chain-type crystal, rhombohedral tellurium. The coefficients of thermal expansion parallel, $\alpha_{\parallel}$, and perpendicular, $\alpha_{\perp}$, to the principal axis for these crystals have the following values at room temperature:

Crystal	$\alpha_{\parallel} \times 10^6$	$\alpha_{\perp} \times 10^6$	$(^{\circ}\mathrm{K})^{-1}$
Graphite	28·2	−1·5	
Tellurium	−1·6	27·2	

The greatest increase in length of the crystal is in the direction of the weakest bonds. In the layer-type crystals this is the direction perpendicular to the planes of the layers. Thus in the graphite crystal and also in white tin and zinc, $\alpha_{\parallel}$ has the largest value. In the chain-type crystals the directions of weakest binding are perpendicular to the direction of the axis of the chains, thus, in the tellurium crystal the largest value is the coefficient $\alpha_{\perp}$. The negative values of the coefficients $\alpha_{\perp}$ in graphite and $\alpha_{\parallel}$ in tellurium can be explained by anisotropy of thermal vibration. In the layer-type crystal the maximum thermal displacements of atoms are perpendicular to the planes of the atomic nets and in the chain-type crystal perpendicular to the axes of the chains.

In the cases considered, the pronounced anisotropy of thermal vibrations can lead to a contraction of the crystal which in the two cases respectively is along the axis of the chain and in directions lying in the plane of the atomic nets.

The changes in dimensions of a quartz crystal on heating are interesting. While the dimensions of the *a* and *c* axes of the unit cell of α-quartz increase on heating, β-quartz contracts on both axes (Fig. 249*a*). It appears that the structures of both modifications of quartz are space lattices of silicon-oxygen tetrahedra linked to each other by inclined valency bonds made of oxygen atoms. The oxygen atoms in the



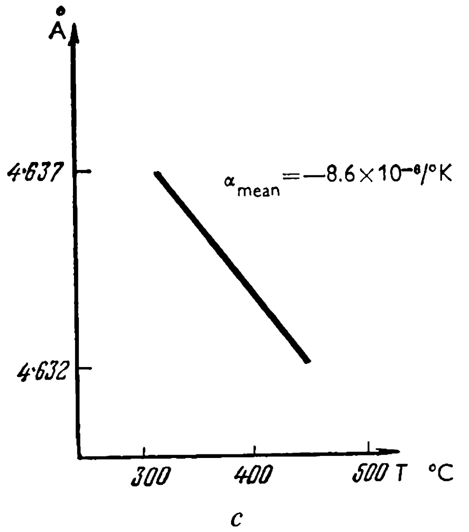


FIG. 249. Thermal expansion: (*a*) curves of the coefficients of thermal expansion of α- and β-quartz (temperature of the $\alpha \rightleftarrows \beta$ phase transition 575° C); the upper curve is the coefficient of linear expansion α_{11} and the lower curve is α_{33}; (*b*) change of volume of the unit cell of lead titanate; the left hand branch of the curve is that of the ferroelectric phase, the right hand branch that of the paraelectric phase (the Curie temperature of the ferroelectric transition T_c = 485° C); (*c*) changes of the lattice parameter of δ-plutonium.

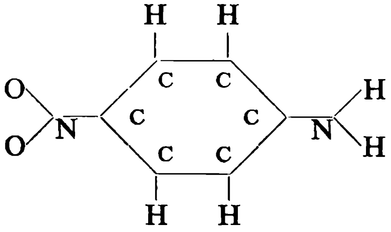
silicon-oxygen nets can execute torsional vibrations similar to those executed by tellurium atoms in the spiral chains of the tellurium crystals. The mechanism of thermal contraction of tellurium and β-quartz may be the same.

The dimensional changes of crystals of ferro- and antiferromagnetics on passing through the Curie point are anomalous. Fig. 249*b* shows the volume change of the unit cell of $PbTiO_3$, which in the ferroelectric range below $T_c = 485°$ C is tetragonal while above the transition point in the paraelectric range it is cubic. Cubic lead titanate behaves like most other crystals, that is it expands on heating, while the tetragonal modification shrinks on heating although the largest contraction takes place at the Curie point. The transition through the ferroelectric Curie point is connected with a sharp change in the specific heat. The reciprocal of the volume of the unit cell of $PbTiO_3$ has a temperature dependence rather like a λ-point.

In ferromagnetic alloys the transition to the paramagnetic state is also accompanied by contraction of the unit cell. If this contraction is comparable in magnitude to the normal thermal expansion then the resultant change of volume may be nearly zero. Good compensation takes place in the alloy of iron with 36% nickel (Invar) which has a zero coefficient of expansion over a wide temperature range.

Very interesting things happen when one of the modifications of plutonium is heated. δ-plutonium which is stable in the temperature range 319-451° C has an f.c.c. structure, that is one of the ordinary simple metallic structures. On heating, the unit cell parameter of δ-plutonium falls almost linearly (Fig. 249*c*), corresponding to a negative coefficient of expansion $\alpha_{\text{mean}} = -8{\cdot}6 \times 10^{-6}/°\text{K}$. It appears that this contraction is connected with the thermal excitation of new valencies. For this reason, the specific volume of the following temperature modification δ'-Pu is equal to the specific volume of δ-Pu, and the specific volume of the highest temperature modification ε-Pu is even somewhat less.

The dependence of thermal expansion on the structure and the strength of the bonds is seen to advantage on the crystal of para-nitroaniline. The molecule of para-nitroaniline has the following structure:



On one side of the benzine ring is the nitro- group NO_2 and on the other side (in the para- position) the amino- group NH_2. As a result of interaction of the molecules between the nitro- group of one molecule and the amino- group of another one, hydrogen bonds may arise, O . . . H – N. This is confirmed by experiment. Fig. 250 shows the projection of the structure of the para-nitroaniline crystal on the monoclinic plane (the hydrogen atoms are not shown). The chain-like distribution of the molecules is easily seen; this is caused by hydrogen bonds which are indicated by dotted lines. The length of the hydrogen bonds is 3·06 Å which is much smaller than the shortest intermolecular spacing 3·66 Å between the chains corresponding to the usual van der Waals' forces. If a polar diagram showing the coefficient of linear thermal expansion *g* of the crystal is drawn for directions which lie in the monoclinic plane, this shows considerable anisotropy of thermal expansion. The smallest value $g = \alpha_{33}$ is approximately in the directions of the hydrogen bonds (in the directions of the molecular chains) and the largest value $g = \alpha_{11}$ in directions approximately perpendicular to the chains. The principal axis of the ellipsoid of thermal expansion $(1+g)^2$ is parallel to α_{11} and makes with the a axis of the crystal an angle of ~40°. The position of the ellipsoid of thermal expansion relative to the crystallographic axes is determined by the peculiarities of the molecular arrangement. Anisotropy of thermal expansion is a result of anisotropy of the intermolecular forces and has its origin in the distribution of hydrogen bonds within the given crystal.

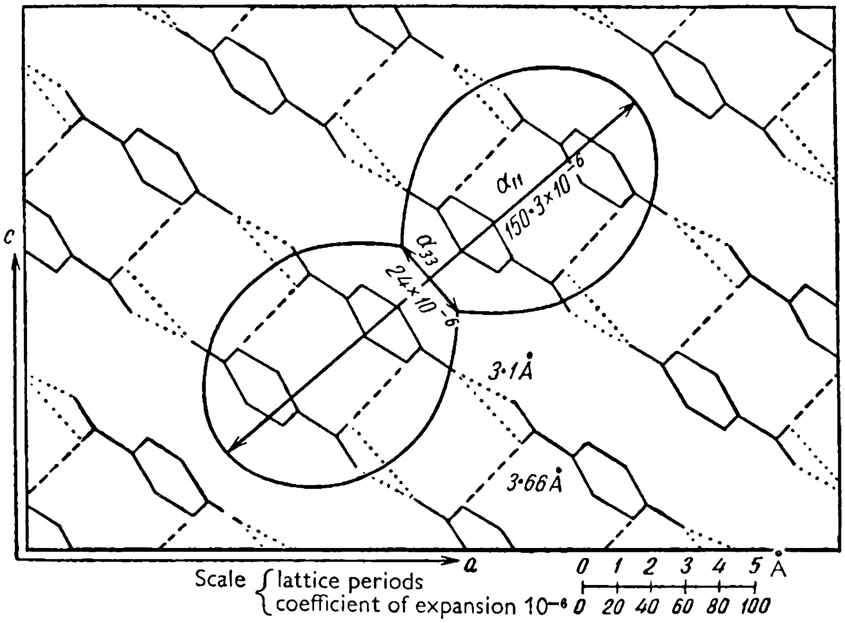


FIG. 250. The connection of the thermal expansion tensor with the structure of the para-nitro-aniline crystal.

Calculation of the Coefficient of Thermal Expansion

At low temperatures atomic vibrations can be considered to be linear, the restoring force to be Hookean and the potential energy of vibration to be parabolic. Any given lattice vibration can be reduced to a system with normal independent harmonic vibrations which do not interfere with each other. If the crystal is homogeneous and there are no stresses such as foreign atoms or mechanical deformations, then elastic waves are propagated in the crystal as in a uniform medium without scattering. The lattice waves will be reflected and scattered only at the crystal boundaries. As the temperature is raised and the amplitude of thermal vibrations increases the situation ceases to be linear. Hooke's law ceases to be obeyed for such vibrations and in the expression for the restoring force there appears a quadratic term

$$f(\xi) = \beta\xi - \gamma\xi^2 \qquad (\beta > 0;\ \gamma > 0), \qquad \ldots [14.110]$$

where ξ is the displacement of an atom from its position of equilibrium. If the coefficient of thermal expansion is positive then the coefficient of anharmonicity is positive. The expression for the potential energy of the oscillator now includes a cubic term

$$U(\xi) = \tfrac{1}{2}\beta\xi^2 - \tfrac{1}{3}\gamma\xi^3. \qquad \ldots [14.111]$$

Consider a linear chain of coupled oscillators. The time average of the number of atoms which have thermally produced displacement ξ is determined by the Boltzmann factor $e^{-U(\xi)/kT}$. Consequently the displacement ξ in a system of oscillators has weight $\xi e^{-U(\xi)/kT}$. The mean displacement is

$$\bar{\xi} = \frac{\int\limits_{-\infty}^{+\infty} \xi e^{-\frac{U(\xi)}{kT}}\, d\xi}{\int\limits_{-\infty}^{+\infty} e^{-\frac{U(\xi)}{kT}}\, d\xi} = \frac{\int\limits_{-\infty}^{+\infty} \xi e^{-\frac{\beta\xi^2}{2kT}} \cdot e^{+\frac{\gamma\xi^3}{3kT}}\, d\xi}{\int\limits_{-\infty}^{+\infty} e^{-\frac{\beta\xi^2}{2kT}} \cdot e^{+\frac{\gamma\xi^3}{3kT}}\, d\xi}. \qquad \ldots [14.112]$$

Expand the anharmonic exponential in series

$$e^{\frac{\gamma\xi^3}{3kT}} = 1 + \frac{\gamma\xi^3}{3kT} + \ldots$$

Assuming that the energy corresponding to the anharmonic restoring force is small compared with thermal energy, $\frac{1}{3}\gamma\xi^3 \ll kT$, then in the denominator of [*112*] we need retain only the first term of the expansion and in the numerator the first two terms. If we had retained only the first term in the numerator then we should have had an odd function of ξ, for which $\bar{\xi} = 0$. This shows that in a system of harmonic oscillators the equilibrium spacing between the particles does not depend on temperature and in consequence $\alpha = 0$.

With the approximations stated above, formula [*112*] takes the form

$$\bar{\xi} = \frac{\gamma \int\limits_{-\infty}^{+\infty} \xi^4 e^{-\frac{\beta\xi^2}{2kT}}\, d\xi}{3kT \int\limits_{-\infty}^{+\infty} e^{-\frac{\beta\xi^2}{2kT}}\, d\xi}. \qquad \ldots [14.113]$$

To integrate this we make use of the formula†

$$\int\limits_{0}^{\infty} x^{2a} e^{-px^2} dx = \frac{1 \cdot 3 \cdot 5 \ldots (2a-1)}{2^{a+1} p^a} \sqrt{\frac{\pi}{p}}. \qquad \ldots [14.114]$$

Carrying out the integration we obtain

$$\bar{\xi} = \frac{\gamma kT}{\beta^2}. \qquad \ldots [14.115]$$

The relative extension of the chain, referred to a single interatomic spacing is

$$\frac{\Delta a}{a} = \frac{\bar{\xi}}{a} = \frac{\gamma kT}{a\beta^2} = \alpha T.$$

---

† See T. B. Dwight: *Tables of Integrals*, Macmillan, New York.

From which the coefficient of linear expansion is

$$\alpha = \frac{\gamma k}{a\beta^2}. \qquad \ldots [14.116]$$

The sign of the expansion coefficient is determined by the sign of γ. If the asymmetry in the potential energy curve is such that when the particles are brought closer together the energy increases faster than when they are removed further apart (this is the case we are considering) then $\gamma > 0$ and consequently heating the body makes it expand.

Express the potential energy of interaction as a Taylor series:

$$U(r) = U(r_0+\xi) = U_0 + \left(\frac{dU}{dr}\right)_0 \xi + \tfrac{1}{2}\left(\frac{d^2U}{dr^2}\right)\xi^2 + \tfrac{1}{6}\left(\frac{d^3U}{dr^3}\right)_0 \xi^3 + \ldots \qquad \ldots [14.117]$$

In equilibrium $\left(\frac{dU}{dr}\right)_0 = 0$. Comparing [111] and [117] we obtain

$$\beta = \left(\frac{d^2U}{dr^2}\right)_0, \quad \gamma = -\tfrac{1}{2}\left(\frac{d^3U}{dr^3}\right) = -\tfrac{1}{2}\left(\frac{d\beta}{dr}\right)_0 \approx \frac{\beta}{2r_0}. \qquad \ldots [14.118]$$

Putting this value of γ into [116] we obtain (for $r_0 = a$)

$$\alpha = \frac{\gamma k}{2a\beta^2} = \frac{k}{2a^2\beta} = \frac{k}{2a^3E} = \frac{N_k}{2Na^3E} = \frac{R}{2vE}, \qquad \ldots [14.119]$$

where E is Young's modulus. If we consider the elastic extension referred to a unit cell we have

$$F = Ea^2\frac{dr}{a} = \beta dr,$$

from which $\beta = Ea$. Putting numerical values in [119] gives

$$\alpha = \frac{R}{2vE} \approx \frac{10^7}{10 \cdot 10^{11}} = 10^{-5}/°\mathrm{K},$$

which, in order of magnitude, agrees with experiment.

Problems

1. (*a*) Find the relationship between coefficients of volume expansion α, volume compressibility β, and thermal elasticity γ; (*b*) verify the formula for the case of an ideal gas.
2. Find the general expression for $c_p - c_v$. (Hint: use the condition for entropy to be a complete differential.)
3. Find the value of $c_p - c_v$ for iron at 700° C [$\alpha = 5\times 10^{-5}/°\mathrm{K}$; $\beta = 5{\cdot}9\times 10^{-7}$ $(\mathrm{kg/cm^2})^{-1}$; density $\rho = 7{\cdot}9$ gm/cm^3].
4. Find an expression for the Einstein specific heat of a solid.
5. (*a*) Find which is the direction of null expansion for a uniaxial crystal, (*b*) discuss the formula you have found and (*c*) find the direction of null expansion for calcite.

6. Find the coefficients of thermal expansion of a uniaxial crystal if the volume coefficient of expansion and the change in the axial ratio c/a on heating are known.
7. Determine the position and magnitude of the axes of the ellipsoid of thermal expansion for a monoclinic crystal.

Chapter 15

SELF DIFFUSION AND DIFFUSION IN CRYSTALS

1. Basic Concepts and Experimental Results

Fluctuation of Thermal Energy. Formation of Vacancies and Interstitial Atoms in a Crystal

Thermal motion of particles in solids (crystals and glasses) consists of vibrations around definite positions of equilibrium, the particles retaining their arrangement in space relative to each other. As a result of fluctuations of thermal energy, some atoms acquire an energy which is much more than the mean. Such atoms may be able to surmount the potential barriers which keep them in equilibrium positions in the lattice and take up non-equilibrium (interstitial) positions between the other atoms. In this case there also appear in the crystalline lattice vacant lattice sites (vacancies) corresponding to equilibrium positions which are not occupied by atoms. A lattice vacancy can readily be occupied by one of the neighbouring atoms; this results in movement of the vacancies through the crystal and in fact in movement of atoms. Interstitial atoms also have higher mobility than the atoms which occupy the equilibrium positions in the lattice. Thus, atoms in a solid can move around at rates which increase with increasing temperature.

Diffusion in solids plays an important part in making possible such processes as thermo-chemical treatment (case-hardening, nitriding, cyaniding, chromating and so on), which produce strengthened and hardened surface layers or protective coatings (oxide films on aluminium surfaces). Diffusion processes play an important part in corrosion and in a variety of phase transitions which take place in solids. Methods of producing metals and alloys from powder (powder metallurgy) by sintering are likewise based on diffusion.

Diffusion transport of atoms proceeds also in pure metals. To distinguish it from chemical diffusion (the transport of chemically dissimilar atoms) the transport of atoms of the same kind is called self-diffusion. Self-diffusion takes place in crystals, liquids and gases. Self-diffusion plays an important part in the processes of recovery and recrystallization which take place when deformed metals are heated.

The Basic Equations of Diffusion

The processes of diffusion are relatively simple if at the beginning there is a non-uniform distribution of the components. When such a specimen is heated, diffusion usually leads to equalising the concentration *c*. If one measures the change of concentration c with time at various points in the specimen one can derive quantities which measure diffusion. In the simplest case the amount of substance dm diffusing in time dt is determined by the equation

$$dm = -DS\frac{dc}{dx}dt, \qquad \ldots [15.1]$$

where $\frac{dc}{dx}$ is the gradient of concentration, S is the cross-sectional area normal to the gradient of concentration through which the diffusion current passes and D is the diffusion coefficient which is a measure of the velocity of diffusion. In the most general case, with a non-uniform distribution of concentration and a non-stationary process such that the concentration varies with time, diffusion is described by a differential equation of the type [*1.13*], in which $v = c$. The equation is then called *Fick's law*. For a stationary process $\left(\frac{dc}{dt} = 0\right)$ equation [*1.13*] goes over to the equation $q = -D\frac{dc}{dx}$ (where q is the diffusion current) corresponding to equation [*1*] above.

The solution of the diffusion equation can be found in various cases. In one such, the smooth surfaces of two pieces of different metals are held together at a temperature sufficiently high for the diffusion rate to be large. The distribution of concentration at each moment of time can be described by curves as shown in Fig. 251. By measuring the distribution of concentration at any instant and comparing the result with the theoretical curve the diffusion coefficients at the temperature of the experiment can be determined. In some cases the rate of diffusion in solids may be very large. Thus at 300° C, the rate of diffusion of gold in lead is greater than that of diffusion of sodium chloride in aqueous solution at room temperature.

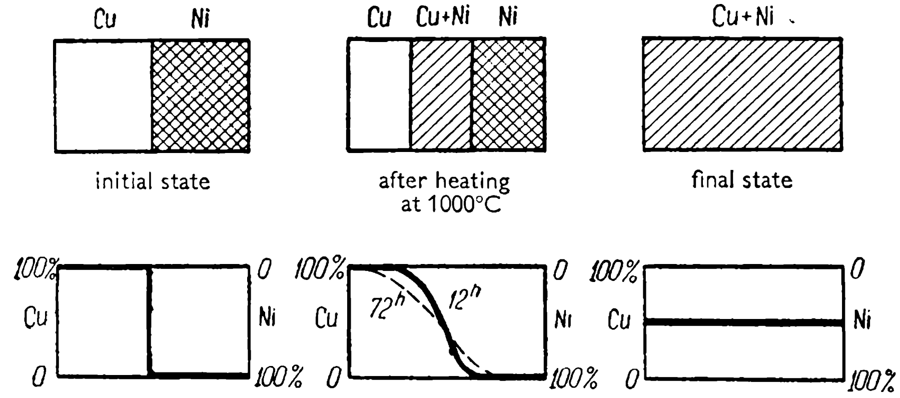


FIG. 251. Different stages of diffusion in a binary alloy.

Let Q be the magnitude of the potential barrier (energy of activation for diffusion), which has to be surmounted by an atom in a crystal if it is to make a diffusion jump. The temperature dependence of the diffusion coefficient is determined by the Boltzmann law

$$D = Ae^{-\frac{Q}{kT}}. \qquad \ldots [15.2]$$

The quantities A and Q are almost independent of temperature but may be highly dependent on composition. In some cases

$$A = Ca^2Q, \qquad \ldots [15.3]$$

where a is the interatomic spacing and C is a constant.

The Measurement of the Self-diffusion Coefficient by means of Tracer Atoms

Formula [*2*] is also applicable to self-diffusion. Up till recently it was not possible directly to measure the rate of self-diffusion experimentally. However since the discovery of artificial radioactivity the position has changed, since it is now possible to measure the rate of mixing of tracer atoms, that is radioactive isotopes of the same element introduced as if they were impurities. The first application of artificial radioactive isotopes for this purpose was carried out in Leningrad by A. Zagrubsky (1937) who measured self-diffusion in gold, the artificial radioactivity being produced by irradiating normal gold with neutrons from a radium-beryllium source. A layer of radioactive gold was deposited electrolytically on the surface of a gold disc. The rate of self-diffusion after heating at various temperatures was determined by measuring the rate at which the activity of the electrodeposited layer fell as active atoms diffused inwards.

The Coefficient of Self-diffusion of Gold

T, °K	1073	1173	1273
D, cm^2/day	$2{\cdot}3 \times 10^{-7}$	$5{\cdot}0 \times 10^{-7}$	$1{\cdot}4 \times 10^{-5}$

The experimental results on the coefficient of self-diffusion of gold lead to the formula

$$D = 7{\cdot}94 \times 10^5 e^{-\frac{62900}{RT}} \text{ cm}^2/\text{day}. \qquad \ldots [15.4]$$

The activation energy for self-diffusion Q is 62·9 kcal/mole and is about 2/3 of the total energy of sublimation of gold (L = 92·0 kcal/mole at room temperature).

Activation Energies for Diffusion and Self-diffusion in Metals

The activation energies for diffusion and self-diffusion in some metals are given in Table 69. The activation energies for diffusion for different atoms in the same solvent may be very different. The greater the chemical difference between the solvent and the diffusing atoms the less is the activation energy; the largest values are for self-diffusion. This generalisation can be explained if we assume that by reducing the chemical difference between the atoms the binding between them is increased.

Carbon and nitrogen form interstitial solid solutions in iron. Since diffusion of these atoms in the iron lattice proceeds fairly rapidly this shows that diffusional movements can take place along interstitial sites. The rate of diffusion depends on the condition of the material (grain size, internal stresses). The rate of diffusion of thorium in tungsten increases markedly as the grain size of the tungsten is reduced. This shows that thorium atoms diffuse preferentially along the crystal boundaries. However diffusion of zinc in copper shows no dependence of rate on grain size; apparently zinc atoms diffuse through the grains as readily as along the grain boundaries. The

Table 69

Activation Energies for Diffusion and Self-diffusion

Solvent	Solute	Q, eV
Cu	Cu	2·1
	Zn (10%)	1·3
	Zn (30%)	1·5
	Sn (10%)	1·1
Pb	Pb	1·2
	Sn	1·1
	Tl	1·1
	Cd	0·65
	Ag	0·65
	Au	0·57
Fe	C	1·2
	N	1·1

rate of diffusion of zinc in copper is highly dependent on the composition of the alloy. As the composition of zinc increases the intermediate phases (β, γ, ε) are formed and along the diffusion path there appears a distribution of the phases (Fig. 252) corresponding to the copper-zinc composition diagram. The rate of diffusion of zinc is different in the different phases.

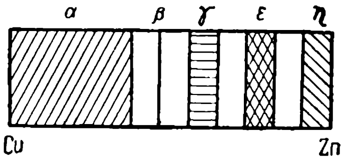


FIG. 252. Distribution of phases in the diffusion path.

2. The Mechanism of Diffusion in Crystals

There are various ways in which atoms can change place in a crystalline lattice. If the crystal has an ideal undamaged structure and is to retain this ideal structure, two

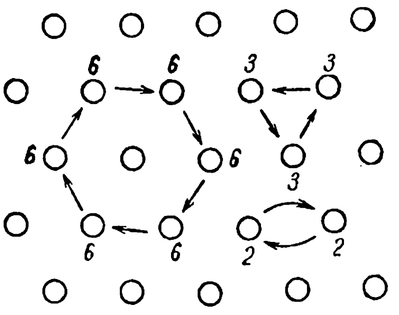


FIG. 253

Diffusion by the mechanism of exchange of sites (exchange of two atoms, three atoms and six atoms).

mechanisms of diffusion are possible (Fig. 253): (*a*) *direct interchange* of a pair of atoms and (*b*) *ring diffusion* when a large group of atoms take part in a simultaneous exchange of their positions. Fig. 253 shows examples of ring diffusion in a triangle and in a hexagon. The latter case corresponds to close packing of spheres, where six atoms form a regular hexagon which approximates to a circle. Direct exchange of positions by pairs of atoms in a close packed structure requires that large potential barriers be surmounted since neighbouring atoms have to pass each other. But group exchanges of atoms by the ring mechanism of diffusion would mean that the share of the energy of the potential barrier belonging to a single atom would be much less than when pairs of atoms change position. However this energy increases as the number of atoms in the ring increases and the probability that this process can take place falls rapidly because of the difficulty of accumulating the necessary fluctuations of thermal energy.

In a real crystal two basic types of structural defect on an atomic scale are possible; vacant lattice sites *A* (vacancies) (Fig. 254) and interstitial atoms *A'*. The probability of diffusion jumps of atoms in a real crystal is greater than in an ideal crystal for two reasons. First, diffusion in a real crystal can take place by the movement of individual atoms, not groups of atoms and secondly, the potential barriers are smaller. Fig. 255*a* shows schematically the curve of the potential energy along one of the atomic rows of a crystal with a strictly periodic structure. The curves of potential energy for a crystal with a defective structure are shown in Fig. 255*b* (defect of the vacancy type) and in Fig. 255*c* (defect of the interstitial type).

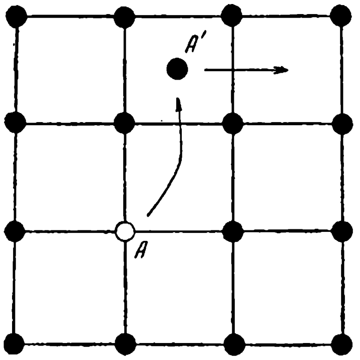


Fig. 254. Simultaneous formation of a vacancy and an interstitial atom.

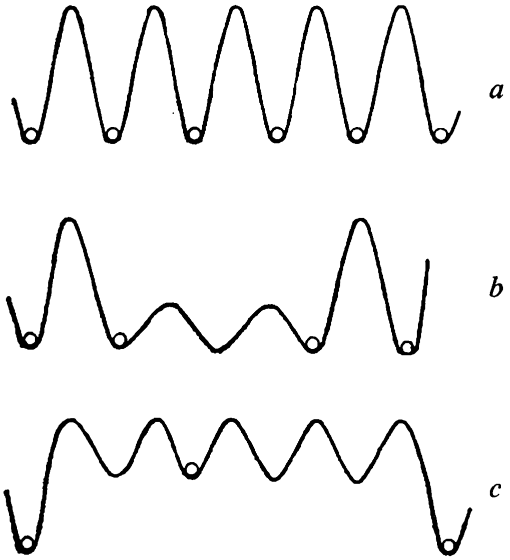


Fig. 255. Potential energy curves for (*a*) a periodic row; (*b*) a vacancy; (*c*) an interstitial atom.

These types of defect structure can arise in an ideal crystal as a result of thermal agitation. Let *u* be the energy necessary to remove an atom from an equilibrium position in the crystal lattice (a lattice site) and put it between the neighbouring atoms (interstitial site) (Fig. 254). This process was considered by Frenkel as a sort of internal 'evaporation' or local dissociation of the crystal. In each process of internal evaporation both a vacancy and an interstitial is formed, so a crystal, even of a simple one component substance, at any temperature different from 0° K is a solid solution of both types of 'impurity': vacancies and displaced atoms.† To each temperature there corresponds a definite equilibrium concentration of the defects of one kind or the other. The number of displaced atoms is proportional to $\exp(-u/kT)$.

Vacancies can also be formed in another way, namely by the removal of one of the atoms of the surface layer (Fig. 256) since such an atom has only partial binding with the crystal and can move to an adsorbed state. The vacancy thus formed can pass into the interior of the crystal by exchanging with neighbouring atoms. Such a process of formation of vacancies can be thought of as dissolving negative atoms in it (Frenkel). Some of the atoms adsorbed on the surface of a crystal can also penetrate inwards and form defects of the second type. If the process of internal evaporation gives equal numbers of defects of both types then the independent processes of formation and destruction of vacancies and interstitials on the surface of the crystal can destroy this equality. The transition of one of the surface atoms to an adsorbed state requires less energy than is necessary for complete evaporation since a smaller number of bonds have to be broken, so this process is more probable than the process of complete evaporation.

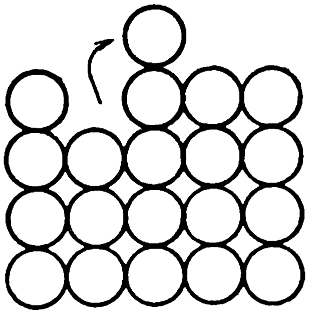


FIG. 256. Formation of a vacancy on the surface of a crystal.

As an example of a calculation of the concentration of atomic defects in a crystal, suppose that $u = L$, where L is the heat of sublimation. In these circumstances the concentration of defects in the crystal is equal to the density of the saturated vapour above the crystal and a crystal containing N atoms has a number N' of defects given by

$$N' = N \exp(-L/RT). \qquad \ldots [15.5]$$

Table 70 gives the numbers of defects calculated by means of formula [5] for a cadmium crystal for which $L = 23$ kcal/mole.

Table 70

Concentration of Atomic Defects in Cadmium

T, °K	N'/N	τ
300	10^{-18}	10^{5} sec $\approx$ 1 day
600	10^{-9}	10^{-4} sec

---

† We may here digress to point out that the majority of crystals of simple substances are isotopic solutions in which various degrees of disorder are possible in the distribution of the isotopes on the crystal lattice. The only exceptions are natural monoisotopic elements, for example sodium, or artificially separated isotopes. Disorder in isotopic mixtures is retained even at absolute zero.

The concentration of defects increases very rapidly with temperature although in the case considered above the final concentration is very small (in some cases the concentration of vacancies can be much larger and, near the melting point, may be as large as 1%). However the absolute number of defects in unit volume of the crystal is very large because the number of the atoms is very large. Thus in one cc. of cadmium there are $N \approx 10^{22}$ atoms. So at room temperature in each cubic centimetre of metal there are $N' = 10^4$ vacancies and near the melting point there are $N' = 10^{13}$. These numbers are similar in order of magnitude to the apparently insignificant numbers of impurities and defects which determine the electrical and optical properties of semiconductors. Here these small concentrations of defects are of basic importance for diffusion in common crystals.

Let us now carry out some calculations on the rate at which atoms move by the vacancy mechanism. First calculate the rate at which vacancies move in the crystal. If Δu is the height of the potential barrier between an atom and a vacancy then $\exp(-\Delta u/kT)$ is proportional to the probability that the atoms which surround a vacancy will 'evaporate' and pass into the vacancy. This probability is inversely proportional to the lifetime of an atom around a vacancy: $\tau = A \exp(\Delta u/kT)$. If there were no potential barriers an atom would move with mean thermal velocity v_0. In this case the lifetime of an atom is equal to the time to move it by one interatomic spacing:

$$A = \tau_0 = \frac{a}{v_0} = \frac{10^{-8}}{5 \times 10^4} = 2 \times 10^{-13} \text{ sec.}$$

This time is of the same order of magnitude as the period of free oscillation of atoms in the lattice. Thus

$$\tau = \tau_0 \exp(\Delta u/kT). \qquad \ldots [15.6]$$

The lifetime of an atom is strongly dependent on temperature. If we use the figures in Table 70, and put $\Delta u = 23$ kcal/mole, then we obtain the lifetimes which are given in the third column. At room temperature a vacancy moves at a rate of one atomic jump per day while at 300° C each vacancy makes 10^4 jumps per second.

We can now calculate the mean rate at which atoms move by the vacancy mechanism. The probability that a given atom has a vacancy beside it is $N'/N = \exp(-u/kT)$. The probability that an atom adjacent to a vacancy will move by one step in time dt is dt/τ. So the probability that a given atom in the crystal will move by one step in the stated time is equal to the probability of the joint event $(dt/\tau)(N'/N) = dt/\theta$, where θ is the lifetime of the atom we are considering. Hence

$$\theta = \tau \frac{N}{N'} = \tau e^{\frac{u}{kT}} = \tau_0 e^{\frac{u+\Delta u}{kT}}. \qquad \ldots [15.7]$$

The mean velocity of atomic movement by the vacancy mechanism is then

$$v = v_0 \exp(-[u+\Delta u]/kT).$$

3. Diffusion in Deformed Alloys

Uphill Diffusion

In deformed alloys which have large internal stresses, diffusion can redistribute the components of the solution in directions contrary to those which lead to uniform distribution. To begin with, consider qualitatively a binary solid solution without internal stresses and with uniform distribution of components A and B. The atomic radii of these components are supposed to be different ($r_A > r_B$). The distribution of the atoms in such a solid solution in statistical disorder is shown schematically in Fig. 257*a*.

Now subject the crystal to non-uniform elastic deformation. For example such a deformation might be elastic bending, in which case those zones which are extended will on the average have an increase in the volume of their unit cells, while those zones which become compressed will have their volume reduced. As a result of the extensions and compressions, the energy of the bent crystal will be greater than the energy of an undeformed crystal by the work to deform it elastically. If the crystal is freed from the external forces which are deforming it, then the internal elastic stresses within it will tend to restore it to the undeformed condition.

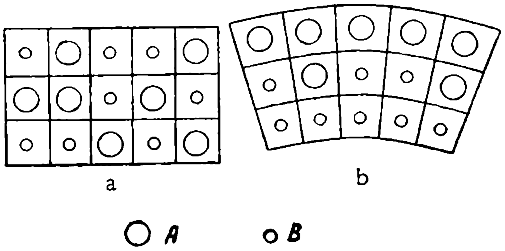


FIG. 257. Diffusion in a solid solution (*a*) without stress; (*b*) with internal stresses.

Suppose that the crystal is held in a bent state and heated to a temperature such that the rate of diffusion is large. Under these conditions, the transition to the more stable state with minimum free energy requires a reduction in the total energy of elastic deformation. In a solid solution where the atoms have different dimensions, reduction in elastic energy can take place if the atoms with larger radius go preferentially into the extended parts, that is the outer zones, while the atoms with smaller radius go to the compressed zones. This is shown in Fig. 257*b*. The possibility of such a diffusion process leading to an increase in the concentration gradient (uphill diffusion) was suggested by Konobeevsky to explain the influence of low temperature annealing on a deformed alloy of magnesium ($r_{Mg} = 1{\cdot}63$ Å) with aluminium ($r_{Al} = 1{\cdot}43$ Å) and zinc ($r_{Zn} = 1{\cdot}33$ Å). In these alloys, annealing produced complex changes in the intensity and width of X-ray lines: as well as the usual $K_{\alpha 1}$ and $K_{\alpha 2}$ doublets such as reappeared when pure deformed magnesium was annealed these alloys in some cases developed 'triplets'. In a special case, diffusion produced an X-ray pattern corresponding to two solid solutions with different aluminium concentrations. This would have resulted in a doubling of the lines on the X-ray pattern, so the original doublets would have become two pairs of lines. The appearance of triplets is explained by partial superposition of these pairs. If after the low temperature anneal, the alloys are heated at higher temperature such that recrystallization can take place, then the redistribution is retained even after recrystallization. But if the deformed alloys are heated at once at high temperature then the X-ray picture of the recrystallized alloy shows normal doublets. This can be explained if in rapid heating at high temperature the internal stresses are quickly released by recovery in the same way as in pure deformed metals. In such short intervals of time, diffusion processes do not take place and consequently uphill diffusion is not observed. The phenomenon of diffusion in deformed alloys shows that it is necessary to generalise Fick's law, the generalised equation having the form

$$\frac{\partial c}{\partial t} = D' \frac{\partial^2 c}{\partial x^2} - D'' \frac{\partial^2 \varepsilon}{\partial x^2}, \qquad \ldots [15.8]$$

where ε is a quantity which measures the energy of elastic stresses, $d\varepsilon/dx$ is the gradient of these stresses, D' the usual diffusion coefficient, and D'' the coefficient of uphill diffusion. The second term in [*8*] has a negative sign since uphill diffusion leads to an increase in the concentration gradient.

4. Diffusion in Solid Electrolytes

Investigations of the rate of diffusion of lead ions in $PbCl_2$ and $PbBr_2$ by means of tracers of artificial radioactive lead, gave the following temperature dependence for the coefficients of diffusion:

$$D_{PbCl_2} = 1{\cdot}06 \times 10^7 e^{-\frac{38120}{RT}} \quad \text{and} \quad D_{PbBr_2} = 3{\cdot}43 \times 10^4 e^{-\frac{30000}{RT}} \text{ (cm}^2\text{/day)}.$$

Hence the rate of diffusion of lead ions in lead bromide is greater than in lead chloride. This is connected with the larger ionic radius of bromine as compared with that of chlorine.

Data on the mobility of ions in solid electrolytes can also be obtained from measurements of the electrical conductivity of crystals. Comparison of the mobilities of ions from diffusion data and electrical conductivity data shows good agreement in the case of lead bromide where the electrical conductivity at high temperatures is due to movement of the cations Pb^{++}. For lead chloride the two sets of data are in serious disagreement: the measured electrical conductivity is about 1000 times larger than the calculated. This shows that in lead chloride the electrical conductivity takes place by the mobility of Cl^- anions.

Investigations at different temperatures show that there is a different temperature dependence of the mobility of cations and anions. As the temperature is lowered the mobility of the lead ions falls faster than that of the iodine ions. At 260° C the electrical conductivity is 40% due to ions Pb^{++} and 60% due to ions I^-, while at 155° C the Pb^{++} ions are responsible for only 0·4% of the electrical conductivity. In these salts at low temperatures mobility is due to anions, and at high temperatures to cations. At low temperatures the electrical conductivity is very sensitive to the mechanical state of the specimen and to impurities.

The electrical conductivity of ionic crystals has a positive temperature coefficient, that is, it increases with increasing temperature but even at high temperatures is still very small for most of them. However a few substances have exceptionally large electrical conductivities which can be explained by peculiarities in their structures. An example of such a substance is silver iodide which exists in three modifications. The structures of the first two modifications γ-AgI (sphalerite lattice) and β-AgI (wurtzite structure) correspond to the well-known low and high temperature modifications of zinc sulphide. The cubic lattice of the high temperature modification α-AgI which is stable above 146° C has a lattice period $a = 5{\cdot}03$ A and contains two particles in the unit cell. The iodine ions form a b.c.c. lattice with random distribution of silver ions on the vacancies of this structure. The density of α-AgI is about

15% higher than β-AgI while the electrical conductivity near the transition temperature is 4000 times greater than that of β-AgI. Near the melting point (552° C) the electrical conductivity of solid α-AgI is somewhat larger than that of molten silver iodide while most other salts experience an increase in electrical conductivity on melting (that of NaCl for example increases by 3000 times). The high electrical conductivity of α-AgI is explained by peculiarities of its structure: α-AgI has a large number of vacancies which can be occupied by silver atoms but normally remain empty since the number of vacancies is greater than the number of silver ions. The activation energy for diffusion of silver ions through these vacancies is very small; $Q = 600$ cal/gm. atom.

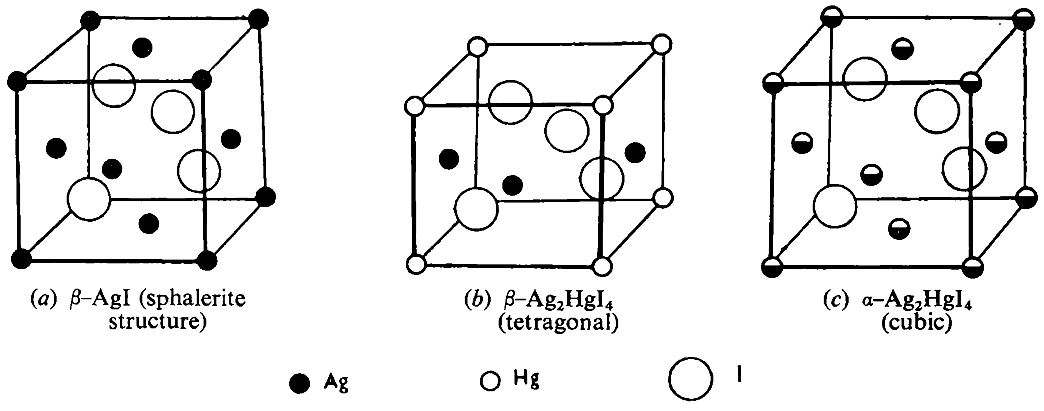

(a) β-AgI (sphalerite structure)
(b) β-Ag_2HgI_4 (tetragonal)
(c) α-Ag_2HgI_4 (cubic)

FIG. 258. Structures of the polymorphic modifications of Ag_2HgI_4.

The hypothesis that there may be permanent vacancies is supported by investigations of the structure and electrical properties of the two polymorphic modifications of Ag_2HgI_4, which have similar structures to β-AgI. Introduction of the divalent atoms of mercury into silver iodide leads to the formation of vacancies according to the scheme $Ag_4I_4 \rightarrow Ag_2Hg\delta I_3$ where the symbol δ indicates that there is a vacancy in the structure. Fig. 258 shows the relation between the structures of these substances. In all cases the iodine atoms form a lattice of the face-centred type. The silver ions in β-AgI occupy the sites of another face-centred lattice (Fig. 258*a*). In β-Ag_2HgI_4 a quarter of the sites of this lattice remain empty (in Fig. 258*b* the centres of the upper and lower faces) and the cell consequently acquires tetragonal symmetry while the shape remains cubic ($c/a = 1{\cdot}00$). The structure of α-Ag_2HgI_4 is cubic with a random

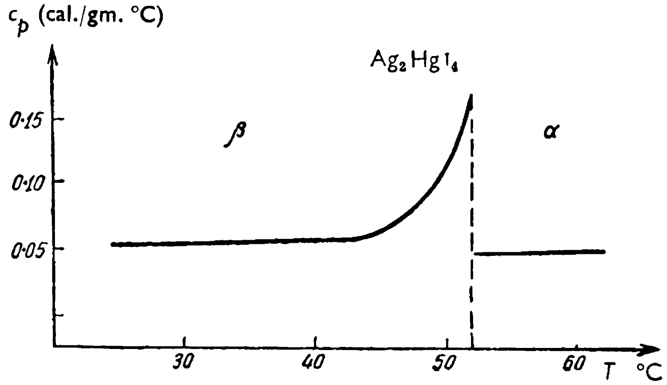


FIG. 259. The change of specific heat at the polymorphic transition β → α Ag_2HgI_4.

distribution of the three cations on the four lattice sites (Fig. 258*c*). The transition from the β- to the α-modification takes place in stages since as the temperature rises the number of cations on the vacant sites increases. This is a second order phase transition from the ordered to the disordered state and is accompanied by a specific heat discontinuity (Fig. 259). The rate of transition increases rapidly, beginning about 40° C and being completed at the Curie temperature $T_c = 50{\cdot}7°$ C. During the transition the colour changes from yellowish to orange.

The electrical conductivity of α-Ag_2HgI_4 is exceptionally large and at the temperatures around 50° C is several thousand times larger than that of any other solid electrolyte. The formula for it is $\sigma = 4\times10^2 \exp(-8600/T)$. Although the ionic radii Ag^+ and Hg^{++} are almost the same, transport experiments show that only 6% of the charge is carried by the mercury ions which have a reduced mobility because of their larger charge. One can think of α-Ag_2HgI_4 as a crystal in which one part of the lattice—the cations—has melted, while the other part—the anion lattice—remains crystalline. In some treatments these crystalline electrolytes are considered as approximating to metals, with the cations taking over the function of the free electrons.

Problems

1. Find the values of Q and A for self-diffusion of gold using Zagrubsky's measurements:

T	800	900	1000° C
D	$1{\cdot}3\times10^{-7}$	$5{\cdot}0\times10^{-7}$	$1{\cdot}4\times10^{-5}$ cm^2/day.

2. The equation for self-diffusion of lead is $D = 5{\cdot}76\times10^5 \exp(-28{,}050/RT)$. (*a*) Find the coefficient of self-diffusion of lead at room temperature, at $T = 165°$ C, and near the melting point; (*b*) at $T = 165°$ C calculate the time which is required for radioactive lead in contact with normal lead to penetrate to a depth of 0·1 mm.

Chapter 16

MECHANICAL PROPERTIES OF CRYSTALS

1. Elasticity, Brittleness and Plasticity of Crystals

The Degree of Deformation

When external forces are applied to a solid, the linear dimensions and shape change in a manner which depends on the type and magnitude of the applied forces. Within the solid, mechanical deformations take place and these can be determined by measuring the dimensions of the specimen before, during and after the application of the stress. In the simplest cases (extension, compression, wire-drawing and rolling) the deformation is determined by the strain

$$\lambda = \frac{l-l_0}{l_0} \cdot 100,$$

where l_0 and l are the initial and final lengths respectively of the specimen. In the case of drawing or rolling the deformation can be described by similar formulae in which the thickness of the sheet or the diameter of the wire is used instead of the length of the specimen. Thus if we use the cross sectional area of the specimen instead of the linear dimensions, we obtain

$$\Lambda = \frac{S_0-S}{S_0} \cdot 100.$$

The strain as measured by Λ, can be transformed to λ, if we assume that volume of a body remains constant in plastic deformation.

Elastic Deformation

External forces change the relative arrangement of the particles in a solid, moving them from the positions of equilibrium which correspond to minima of potential energy. The resulting interatomic interactions give rise to forces within the body which tend to return the particles to their initial positions. These internal forces act against the external forces and when a stationary state is reached are equal to them. If the external forces do not exceed some limit which varies within a wide range for crystals of different substances, then the movement of the particles and the deformation produced are reversible. Such deformations are called elastic. When the external forces are removed an elastically deformed crystal returns to its initial state under the influence of the internal forces. In an elastically deformed state, the magnitude and character of the deformation can be determined from changes in the external shape of the crystal. The magnitude of the internal elastic stresses can be determined from the magnitude of the applied stresses acting on the crystal. The energy of elastic deformation stored in the crystal can be measured by determining the work done by the external forces. Knowing the elastic constants of materials makes it possible to calculate these quantities for various cases of deformation.

Brittleness and Plasticity

When the external forces are increased beyond the limit of elastic reversible deformation, two things may happen—*brittle fracture* or *plastic deformation*—depending on the nature of the crystal and the conditions of the experiment. In fracture the crystal disintegrates irreversibly into two or more parts. As a result of anisotropy brittle fracture of crystals in the majority of cases proceeds along well defined crystal planes which are characteristic of the given crystalline substance. These are the planes which have minimum resistance to fracture and are called *cleavage planes.* The magnitude of the *tensile strength* is an important characteristic of a material.

In many substances there is before fracture a more or less extensive region of irreversible *plastic* deformation, in which the body changes its shape but is not broken in pieces. Examples of plastic crystalline substances are many metals and alloys. Single crystals of zinc can be extended to ten times their original length without fracture. In such a large deformation, a cylindrical single crystalline specimen is transformed into a ribbon.

Plastic deformation leads to marked changes of the structure and properties of crystalline substances. However even at the largest strains the crystalline state and the crystal type are retained.

Plastically deformed crystals exhibit a variety of characteristic forms of damage to their crystalline structure accompanied by the accumulation of the energy of *residual* internal stresses. These strains and stresses balance each other inside deformed crystals and remain after the external forces have been removed. A plastically deformed crystal is in a state of metastable equilibrium. Mechanical, thermal, electrical, magnetic and other properties of deformed crystals may be very different from the properties of undeformed crystals.

Investigations of plastic deformation of crystalline substances and the structural damage produced is of great practical value. The characteristics of plastic fracture in tension are usually, the extension to fracture and the tensile strength, both referred to the initial cross-sectional area of the specimen. In extension the cross-sectional area is reduced and in the region of plastic fracture often shows considerable thinning ('necking'). Thus the true *fracture stress* may be very much greater than the tensile strength. For ideal elastic substances the residual extension is zero.

The Influence of the Conditions of the Experiment on the Type of Deformation

It is impossible to draw a firm line between plastic and brittle crystals. One and the same solid may exhibit brittleness or plasticity. The character of the deformation is influenced by various factors, such as temperature, the type of stress, the rate of deformation, the surrounding medium and so on. Raising the temperature as a rule brings about increased plasticity, this being widely used in hot working of metals by pressure. As the temperature is lowered, brittleness increases. Thus it is usual to obtain powders from plastic substances for X-ray study by filing them in liquid air. The opposite problem—that of retaining the plasticity of materials at low temperatures—arises in a variety of technical problems (cold setting of plastics). In cryogenic technology materials are necessary which retain their plasticity at temperatures down to absolute zero.

The only materials which retain significant plasticity under these conditions are some metals and alloys. Consequently, this is one of the basic properties of metallic binding and is a result of the electron gas in a metal which distinguishes metallic from other types of binding (ionic and covalent). The increasing plasticity with rise in temperature can lead to a peculiar type of deformation which occurs when cadmium crystals are extended at high temperatures (Fig. 260). Because there is no work hardening, slip is concentrated on a few planes. Deformation becomes non-uniform, the crystal is divided into parallel sided blocks and this leads to a lowering of the total plasticity of the crystal. Similar things happen when tin crystals are extended slowly (Fig. 261).

The rate of deformation has a marked influence on the character of deformation. Raising the rate of deformation increases the tendency to brittle fracture. This is true both for crystals and for liquids. With viscous liquids, it is possible to select velocities of deformation at which the type of deformation corresponds to brittle fracture. Thus Fig. 262 shows plastic deformation of a viscous filament made of a mixture of transformer oil and resin (viscosity 5×10^3 poise, which corresponds to the viscosity of honey or treacle). This experiment was carried out by Kornfeld and Rivkin who found that when the filament was pushed slowly it bent plastically (Fig. 262) but when the velocity of the blow exceeded 23 m/sec, the filament broke up into pieces (Fig. 263).

The influence of a stressed condition on the character of deformation is shown by experiments with brittle materials. Thus for example uniaxial stressing of marble (compression or extension) shows it to be brittle, but when it is deformed under conditions of combined volume and tensile stress it acquires some plasticity. Finally, the influence of the medium is shown by experiments on the bending of rock salt crystals. At room temperature in air these crystals are brittle while when they are deformed in water they become plastic.

2. Elasticity of Crystals

Hooke's Law

The basic types of deformation of solids, to which all other types of deformation can be reduced, are: (1) linear extension and compression (Fig. 264*a*) and (2) simple shear (Fig. 264*b*). Corresponding to this, elastic properties of isotropic solids are determined by two constants; Young's modulus E and the shear modulus G. Poisson's ratio μ which measures the transverse compression during extension can be expressed in terms of these two constants. The basic law of plastic defoimation is linear dependence between the forcc F and the deformation Δl produced by it. In uniaxial extension deformation is described by Hooke's Law

$$F = ES\frac{\Delta l}{l} \quad \text{or} \quad p = E\lambda, \qquad \ldots [16.1]$$

where l and S are the initial length and cross-sectional area of the specimen, p is the

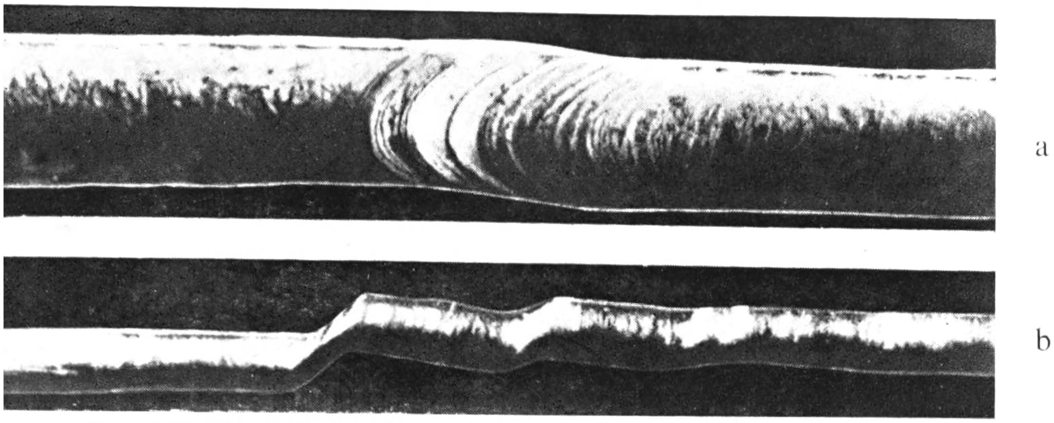


FIG. 260. Formation of slip packets during the extension of a cadmium crystal (a) at 458°K; (b) at 523°K.

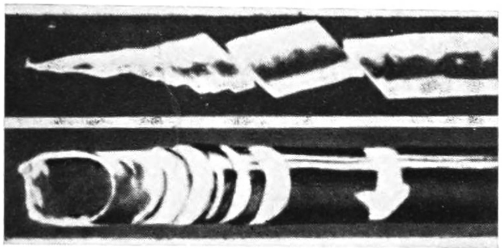


FIG. 261. Formation of slip packets during slow extension of a tin crystal.

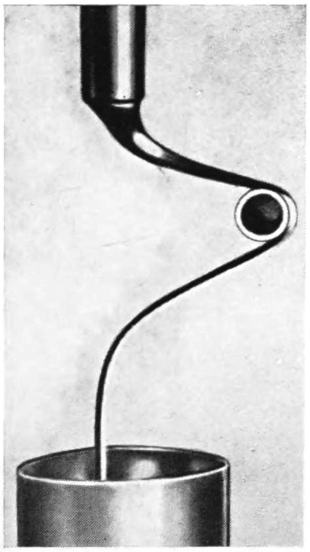


FIG. 262. Plasticity of a filament of a viscous liquid under conditions of slow deformation.

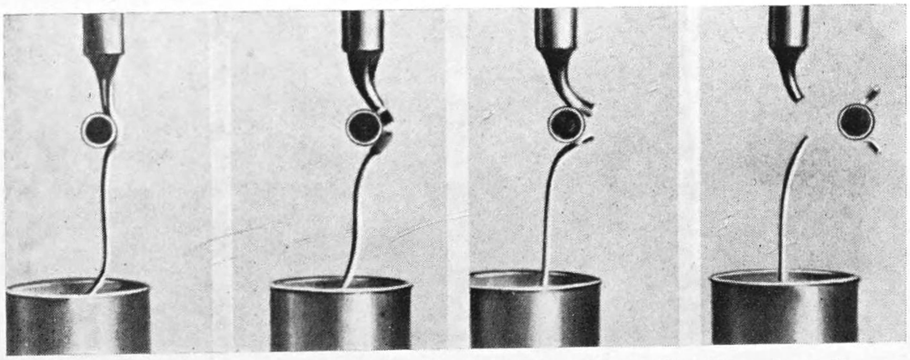


FIG. 263. Brittle fracture of a filament of a viscous liquid as a result of impact deformation.

tensile stress and λ the relative deformation. In simple shear, the relation between the force and deformation is given by a similar relation

$$F = GS\frac{\Delta l}{h} = GS\tan\alpha \quad \text{or} \quad \sigma = G\tan\alpha, \qquad \ldots [16.2]$$

where S is the cross-sectional area of the specimen on the shear plane, h is the thickness in a direction perpendicular to the shear plane, α is the shear angle and σ is the shear stress. In the case of triaxial compression

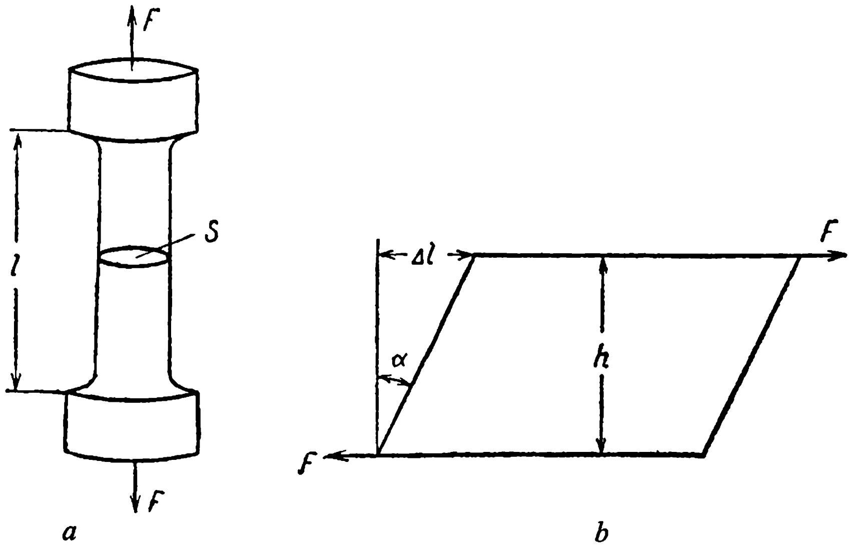


FIG. 264. Basic types of deformation: (*a*) linear extension; (*b*) simple shear.

$$p = -\kappa\frac{\Delta v}{v} = -\kappa\Omega, \qquad \ldots [16.3]$$

where Ω is the volume deformation and κ is the modulus of compressibility.

The linear relation between force and elastic deformation for solids is explained by the fact that the curve of potential energy of interaction between the particles can be approximated by a parabola at sufficiently small displacements. This corresponds to the model of a solid in the form of a system of quasielastic oscillators. If the elastic limit (the onset of plastic flow) is sufficiently high, then departures are possible from the linear Hooke's law on account of the influence of anharmonic terms in the displacement potential energy.

For crystals and other anisotropic media, the relation between stress and deformation has a more complex form than that described by the formulae [*1*] and [*2*]. In such cases the values of the moduli E and G depend on the orientation of the specimen relative to the crystallographic axis. The corresponding number of elastic constants for a crystal is greater than for an isotropic body and changes depending on the symmetry of the crystal.

The Stress Tensor

To determine the characteristics of the stressed state, we subject a small cube to stress and determine components of the forces acting on its faces. When the forces are in equilibrium, this means that they are equal on opposite faces so it is sufficient to investigate the forces acting on three mutually perpendicular faces. Call these forces X, Y, Z, and their components on the coordinate axis x_x, y_x, z_x and so on (Fig. 265). Altogether we have nine components of stress which determine a tensor of second rank:

$$\begin{bmatrix} x_x\, y_x\, z_x \\ x_y\, y_y\, z_y \\ x_z\, y_z\, z_z \end{bmatrix} \qquad \ldots [16.4]$$

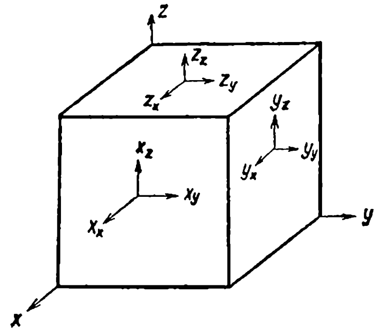


FIG. 265. Resolution of the forces acting on the faces of an elementary cube.

The diagonal components x_x, y_y, z_z are perpendicular to the three faces of the cube and are tensile or compressive stresses. The other six components are parallel to the faces of the cube and are shear stresses. These components form pairs of forces tending to turn the cube round the coordinate axes. In equilibrium, the moments of corresponding pairs are equal $x_y = y_x$, $y_z = z_y$, $z_x = x_z$. The stress tensor is thus symmetrical and is determined by six components, that is, it is obviously represented by a surface which in the general case is a three-axed ellipsoid.

We may write the stress tensor in the form

$$T_{\text{stress}} = \begin{bmatrix} t_{11}t_{12}t_{13} \\ t_{21}t_{22}t_{23} \\ t_{31}t_{32}t_{33} \end{bmatrix}, \quad (t_{ik} = t_{ki}). \qquad \ldots [16.5]$$

The Deformation Tensor

As a result of deformation the relative positions of particles of the crystal are changed. Let R be the radius vector and (xyz) the coordinates of a point before displacement, while R' and $x'y'z'$ are the corresponding quantities after displacement (Fig. 266); u is the displacement vector and u_x, u_y, u_z are its components. If the components of the displacement vector are constant for all points in the body then

$$u_x = x'-x = \text{const}; \quad u_y = y'-y = \text{const}; \quad u_z = z'-z = \text{const},$$

which corresponds to parallel transport of the whole crystal. In deformation, however, the displacements of different points are different. Consider how homogeneous deformation, that is deformation in which a given straight line in the body remains straight and a plane remains a plane. In these conditions a cube is deformed into a parallelepiped and a sphere into an ellipsoid. Such conditions are not satisfied by such deformations as bending, torsion etc. In homogeneous deformation, the components of the displacement vector are linear functions of the coordinates

$$\begin{aligned} u_x &= r_{11}x+r_{12}y+r_{13}z, \\ u_y &= r_{21}x+r_{22}y+r_{23}z, \\ u_z &= r_{31}x+r_{32}y+r_{33}z. \end{aligned} \qquad \ldots [16.6]$$

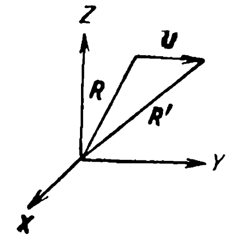


FIG. 266. Displacement of points of a body as a result of deformation.

The nine quantities r_{ik} determine the components of the deformation tensor which is also a tensor of second rank. The components r_{11}, r_{22}, r_{33} are the relative extensions along the coordinate axis; the components r_{23}, r_{31}, r_{12} are shears. The possibility of representing uniform deformation by means of an ellipsoid shows that, in the absence of rotation, the deformation tensor is symmetrical with six independent components. We can then write the deformation tensor in the form

$$T_{\text{def}} = \begin{bmatrix} r_{11}r_{12}r_{13} \\ r_{21}r_{22}r_{23} \\ r_{31}r_{32}r_{33} \end{bmatrix}, \quad (r_{ik} = r_{ki}). \qquad \ldots [16.7]$$

The Elasticity Tensor

Hooke's law for a crystal implies a linear dependence between the components of the stress tensor and the components of the deformation tensor and is represented by the elasticity tensor which is a tensor of fourth rank; $T_{\text{stress}} = cT_{\text{def}}$. In the general case a tensor of fourth rank is determined by a matrix with $9 \times 9 = 81$ terms, while for symmetric tensors of second rank, the arbitrary tensor of fourth rank has $6 \times 6 = 36$ components. We can write the elasticity tensor in the form

$$\begin{array}{c|cccccc} & r_{11} & r_{22} & r_{33} & r_{23} & r_{31} & r_{12} \\ \hline t_{11} & c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ t_{22} & c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ t_{33} & c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} \\ t_{23} & c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} \\ t_{31} & c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} \\ t_{12} & c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} \end{array} \qquad \ldots [16.8]$$

or in abbreviated form

$$t_{ik} = c_{il} r_{lm}. \qquad \ldots [16.9]$$

Solving the linear relations [9] for deformation, we obtain

$$r_{ik} = s_{il} t_{lm}. \qquad \ldots [16.10]$$

The coefficients of proportionality c_{il} are called the elastic constants and s_{il} the elastic moduli.

Consider now the expression for the energy of elastic deformation. For an isotropic body Hooke's law is $t = cr$ and the elastic energy $-dA = tdr$. For an anisotropic body the generalised expression for the elastic energy has the form

$$-dA = t_{11}dr_{11} + t_{22}dr_{22} + t_{33}dr_{33} + t_{23}dr_{23} + t_{31}dr_{31} + t_{12}dr_{12} \qquad \ldots [16.11]$$

The elastic energy of a crystal is a single valued function of the deformation r_{ik}, and the differential of the elastic energy which can be written in the form

$$dA = \frac{\partial A}{\partial r_{11}} dr_{11} + \frac{\partial A}{\partial r_{22}} dr_{22} + \ldots + \frac{\partial A}{\partial r_{12}} dr_{12}, \qquad \ldots [16.12]$$

is a complete differential. Comparing [11] and [12] we obtain

$$t_{ik} = -\frac{\partial A}{\partial r_{ik}}. \qquad \ldots [16.13]$$

Differentiating this expression again with respect to r_{lm} we obtain

$$\frac{\partial t_{ik}}{\partial r_{lm}} = -\frac{\partial^2 A}{\partial r_{ik}\partial r_{lm}} = \frac{\partial t_{lm}}{\partial r_{ik}}, \qquad \ldots [16.14]$$

and since

$$c_{il} = \frac{\partial t_{ik}}{\partial r_{lm}} \text{ and } c_{li} = \frac{\partial t_{lm}}{\partial r_{ik}},$$

it follows that

$$c_{il} = c_{li}. \qquad \ldots [16.15]$$

Consequently, the elasticity tensor is symmetric, being determined in the general case by 21 independent components. It follows that the total number of elastic constants for a triclinic crystal is 21.

The number of elastic constants is reduced depending on the symmetry of the crystal (Table 71) and in the cubic crystal falls to three. As Table 71 shows the number of different symmetry groups of the elastic tensor, including the isotropic group, is 10.

Table 71

The Number of Elastic Constants for Crystals of Different Systems

System	Symmetry	Number
Triclinic	1	21
Monoclinic	2/*m*	13
Rhombic	*mmm*	9
Rhombohedral	3	7
Rhombohedral	3*m*	6
Tetragonal	4	7
Tetragonal	4/*mmm*	6
Hexagonal	6/*mmm*	5
Cubic	*m*3*m*	3
Isotropic		2

The Cauchy Relations

If the forces of interaction between the particles of a crystal have spherical symmetry, then this gives six further relations between the elastic coefficients. These are called the Cauchy relations:

$$\begin{aligned} c_{23} &= c_{44}; \quad c_{56} = c_{14}; \quad c_{64} = c_{25}, \\ c_{31} &= c_{55}; \quad c_{12} = c_{66}; \quad c_{45} = c_{36}. \end{aligned} \qquad \ldots [16.16]$$

For a crystal of the cubic system, formulae [*16*] give one independent relation

$$c_{12} = c_{44}. \qquad \ldots [16.17]$$

Table 72 gives values of the elastic constants for several cubic crystals (non-metals).

The Cauchy relations are approximately satisfied for ionic crystals with structures of the NaCl type, but for AgCl there is a considerable departure from this relation. This also happens for MgO. This can be explained by a strengthening of the directionality of binding in these crystals. The right-hand part of Table 72 gives data for

Table 72

Elastic Constants for Cubic Ionic Crystals

($\times 10^{11}$ dyne/cm^2)

Crystal	c_{11}	c_{12}	c_{44}	Crystal	c_{11}	c_{12}	c_{44}
NaCl	4·94	1·37	1·28	AgCl	6·05	3·64	0·624
NaBr	3·30	1·31	1·33	MgO (80° K)		8·56	15·67
KCl	3·70	0·81	0·79	CaF_2	16·4	4·48	3·38
KBr	3·30	0·58	0·52	FeS_2	36·1	−4·74	10·55
KI	2·67	0·43	0·42	$NaClO_4$	6·5	−2·1	1·20

several crystals in which there are no centres of symmetry. Here the Cauchy relations are not obeyed. The departure from the relation is smallest for fluorite which is nearest to the ionic type.

It is of interest to give the corresponding data for metals. Table 73 gives values of the elastic constants for several face-centred cubic metals and also for tungsten which is body-centred cubic.

Table 73

Elastic Constants for Cubic Metals

$c \times 10^{11}$ dyne/cm^2

Crystal	c_{11}	c_{42}	c_{44}	μ
Cu	17·0	12·3	7·52	0·34
Ag	12·0	8·97	4·36	0·37
Au	19·4	16·6	4·0	0·42
Al	10·8	6·2	2·8	0·34
W	51·3	20·6	15·3	0·17

Metals which have the most symmetrical structures still do not obey the Cauchy relations. Metallic binding cannot be thought of even approximately as spherically symmetric.

The data on the elastic constants of silver chloride (Table 72) are interesting for this has a structure of the NaCl type but relation [*17*] is not obeyed even approximately. This shows that the silver ion does not have spherical symmetry: its electron configuration is different from the configuration of the atom of a noble gas. Judged by its elastic anisotropy, the silver chloride crystal approximates to that of a metal. It is sometimes called a 'transparent metal' and is used in photo-elastic work to make transparent models which reproduce the mechanical properties of polycrystalline metals more accurately than perspex models.

Generalisation of the Theory of Elasticity of Crystals

One of the basic assumptions of the classical Cauchy theory of elasticity of crystals is the symmetry of the stress and deformation tensors. Laval (1951) showed however that when non-central forces act between the particles of an anisotropic lattice the symmetry assumption for these tensors may not be satisfied. The breakdown of the relations $t_{ik} = t_{ki}$ and $r_{ik} = r_{ki}$ is to be expected in crystals with directed bonds without a centre of symmetry and which have enantiomorphous forms, that is, they have either right-handed or left-handed screw axes (for example, quartz, selenium, tellurium etc.). The assumption about the symmetry of deformation and stress tensors is valid for crystals which have no polar directions or have a centre of symmetry, and for these crystals the Cauchy theory of elasticity applies. However this theory must be considered as a special case of a more general theory of the elasticity of crystals.

As was shown above, the elasticity tensor is in the general case determined by a matrix of fourth rank and having 81 terms. However, in the general case relations [*15*] are valid for the elasticity tensor and these require it to be symmetrical relative to the principal diagonal. Consequently in the generalised theory the maximum number of elastic constants is 45 while in the Cauchy theory it is 21.

The new theory of elasticity still requires more complete experimental justification but there is a variety of results to show its usefulness. Thus, in a tetragonal crystal the velocities of propagation of transverse elastic waves along the tetragonal [001] axis with direction of vibration along [100] is

$$v_{[001]} = \sqrt{\frac{N_{44}}{\rho}}$$

while along the [110] axis with direction of vibration along the [001] it is

$$v_{[100]} = \sqrt{\frac{N_{55}}{\rho}},$$

where ρ is the density of the crystal. In the Cauchy theory $c_{44} = c_{55}$ and both velocities should be the same. The measurements of Le Corre on ammonium phosphate show that these velocities are different and that the ratio of the coefficients is quite different from unity. In fact

$$\frac{N_{44}}{N_{55}} = 1{\cdot}36 \pm 0{\cdot}11.$$

Joel and Wooster carefully examined the results of the measurement of the elastic constants of ammonium phosphate obtained on the Laval theory by Zwikker, and found a smaller value of the ratio $N_{44}/N_{55} = 1{\cdot}08 \pm 0{\cdot}01$. This departure is much larger than the probable error and cannot be explained on the Cauchy theory.

Consideration of the piezoelectric properties of quartz from the point of view of the new theory showed that in addition to the two constants d_{11} and d_{14} known before, a third piezo-electric constant d_{39} was now required. The meaning of this constant is that when the quartz crystal is bent round an optic axis which coincides with the direction of the screw axes, piezoelectric charges must be expected to arise on the basal planes. The magnitude of these charges is

$$q = \frac{d_{39}P}{2l},$$

where P is the moment of the couple and l is the thickness of the platelet. In Le Corre's experiments on quartz, a piezoelectric effect was obtained in torsion and this gave

$$d_{39} = -7{\cdot}2 \pm 0{\cdot}7 \times 10^{-8} \text{ c.g.s. units},$$

which is comparable with the magnitude of the other piezoelectric moduli of quartz.

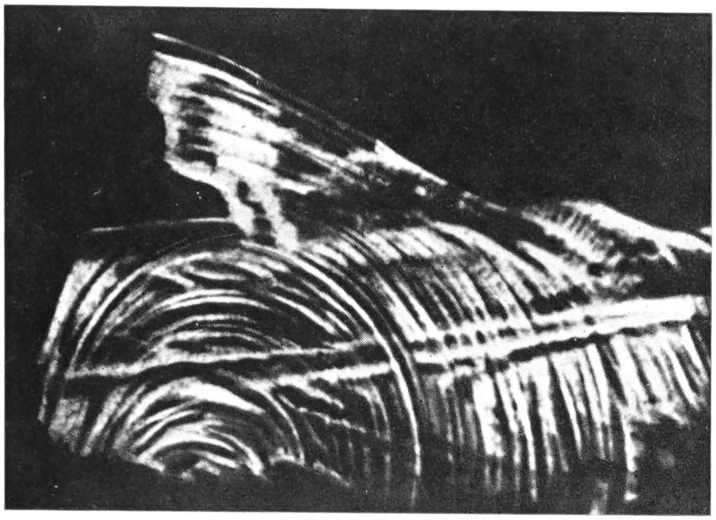


FIG. 267. Appearance of the surface after impact fracture.

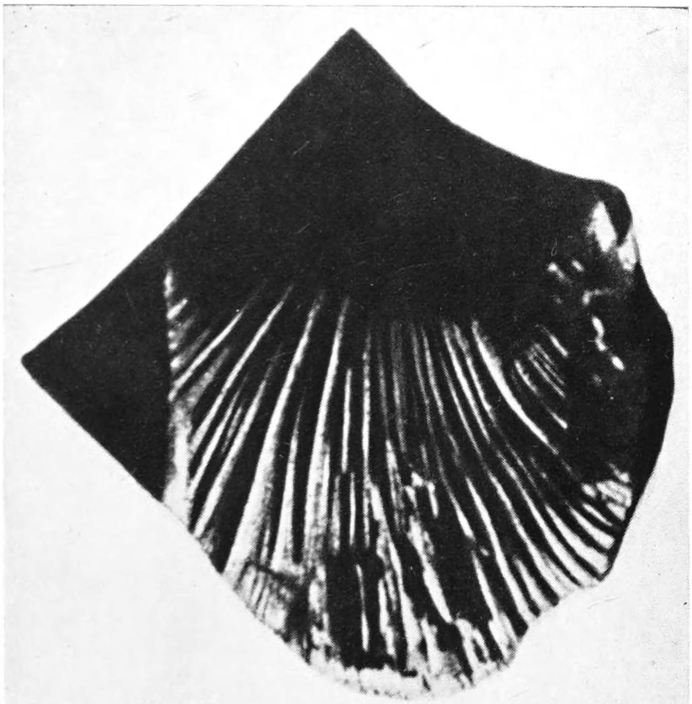


FIG. 268. 'Fracture Rose'.

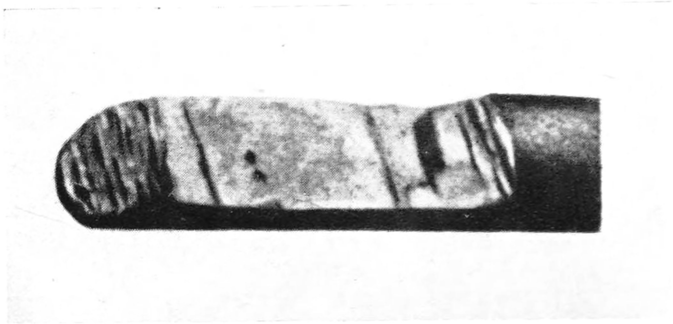


FIG. 269. Cleavage surface of a crystal of tellurium.

Zubov and Firsov measured the plastic constants of quartz by a dynamic method and again showed that the experimental data agreed better with the Laval theory than with the Cauchy theory.

The importance of the new theory of elasticity in the physics of crystals is that it is relevant to all the physical properties of crystals which are connected with stress and deformation (piezoelectricity, magnetostriction and photoelasticity). The Tables of constants which determine the corresponding properties, can be extended to include additional constants required by the new theory. For those crystals where the anisotropy of binding and of structure are small—and up to now, one has to deal mainly with such crystals—the new theory of elasticity agrees with the earlier.

3. Cleavage and the Structure of Crystals

Cleavage of Crystals

When amorphous bodies, such as glass and resin, are broken, one commonly observes that the fracture surfaces are curved into conchoidal shape. When fracture is the result of impact, finer wavy details of the surface relief may arise in the form of concentric circles (Fig. 267), or of systems of lines radiating fanwise from the point of impact. This is shown in Fig. 268 and was called a 'fracture rose' by Gogoberidze. Conchoidal fracture also takes place in some crystals (quartz, corundum, and others). However, more typical of the brittle fracture of crystals is the appearance of plane fracture surfaces called cleavage surfaces. The phenomenon of cleavage and the formation of plane surfaces of separation is an immediate consequence of the anisotropy of the crystal and the structure of its space lattice. The strength of the binding of the various systems of atomic networks to each other is not the same throughout the crystal. The cleavage planes are the planes of weakest binding of the particles within the crystal.

As is well known, crystals of rock salt can be cleaved on the cube $\{100\}$ planes which are the usual external faces. It is not however necessary that the cleavage planes should be the external faces. Thus in crystals of fluorite CaF_2 which crystallize in the form of cubes, the cleavage faces are the octahedral planes $\{111\}$. Cleavage faces are also found in the fracture of some metallic crystals, in particular metals with structures differing from simple packing of spheres. A typical example is tellurium, a cleavage surface of which is shown in Fig. 269.

The Tensile Strength

The strength of a crystal on any given plane cross-section which coincides with one of the crystallographic planes is determined by the *tensile strength* or *fracture stress* v_{hkl}. This quantity is different for different crystallographic planes (*hkl*) and is least for the cleavage planes. For crystals of rock salt, zinc, antimony and tellurium it is $\sim(0{\cdot}5-1{\cdot}0)$ kg/mm^2. The magnitude of the tensile stresses which are necessary to produce fracture of a crystal depends on the angle between the cleavage plane and the direction of the stress

$$F = \frac{\nu S_0}{\cos^2 \phi}. \qquad \ldots [16.18]$$

Here ϕ is the angle between the normal to the cleavage plane and the direction of the Force F and S_0 is the area of the cross-section of the specimen, here assumed to be regular.

The tensile strength on a given cross-section of a crystal is proportional to the total number of bonds per unit cross-section which have to be broken to break the crystal into two parts. If the bonds are not all the same, the strength of the crystal, assuming that all bonds are to be broken simultaneously (the ideal case) is determined by the sum of the contributions of all the bonds. If the crystal is composed of particles all of the same kind, then the fracture stress v_{hkl} on any plane falls as the interplanar spacing d_{hkl} is increased. In this case, the cleavage planes will be the planes with the largest interplanar spacing, that is, the planes which are most densely packed with atoms. This rule, first enunciated by Bravais, is found to be true in some cases (the layer type crystals and close packed metals) but is not true in general. It is not correct, for example, for symmetrical coordination crystals (e.g. sphalerite) which are made up of particles of different kinds since it takes no account of the distribution of particles of the various kinds on the different planes. A more complete analysis will show at least for the simplest cases, the connection between crystal structure and cleavage.

Crystals with Chain-type Structures

By the very nature of chain-type crystals, the interaction between the atomic chains is much less than the interaction within the chains. The strength of a crystal made up of densely packed chains is large in directions parallel to the chains and small in the perpendicular directions. The cleavage planes of a chain-type crystal must be parallel to the axes of the chains, and since dense packing of chains usually requires hexagonal or trigonal symmetry the cleavage planes are prism planes $\{10\bar{1}0\}$. This is supported by the cleavage of crystals of trigonal γ-selenium and tellurium and also by cleavage of asbestos and other fibrous minerals (pyroxenes and amphiboles) which contain chains and ribbons formed of SiO_4 tetrahedra connected by their apexes.

Crystals with Layer-type Structures

In layer-type crystals, interatomic bonds within the layers are stronger than the binding between the layers. Such crystals will readily exhibit cleavage and the cleavage planes are parallel to the layers. Thus, in hexagonal layer-type crystals the cleavage planes are parallel to the basal planes (0001). In agreement with this is the cleavage of the isomorphic crystals arsenic, antimony and bismuth and also of the crystals of zinc and cadmium which have hexagonal close-packed lattices somewhat extended along the hexagonal axis. The layer-type crystals of graphite and boron nitride which have similar structures also cleave perfectly on their basal planes. The three- and two-layer crystals of cadmium chloride and iodide $CdCl_2$ and CdI_2 are constructed of alternate layers of full and empty octahedra (Fig. 270*a* and *b*), while the crystal of molybdenite MoS_2 is composed of alternating layers of full and empty trigonal prisms (Fig. 270*c*). These crystals, just like the layered monoclinic crystals of mica and gypsum $CaSO_4 . 2H_2O$, have well defined cleavage along the lattice layers.

Crystals with Three-dimensional Coordination

If the bonds are uniformly distributed within the space lattice of the crystal then there may be no cleavage planes and the fracture surface may be conchoidal as in quartz α-SiO_2 and corundum α-Al_2O_3.

Crystals of complex compounds such as $BaSO_4$, $K_4Fe(CN)_6.3H_2O$ and others are made up of structural elements of different kinds. The simple ions K^+, Rb^+, Ba^{++}, Cl^- have spherical symmetry. The complex ions SO_4^{-2}, $Fe(CN)_6^{-4}$, SCN^-

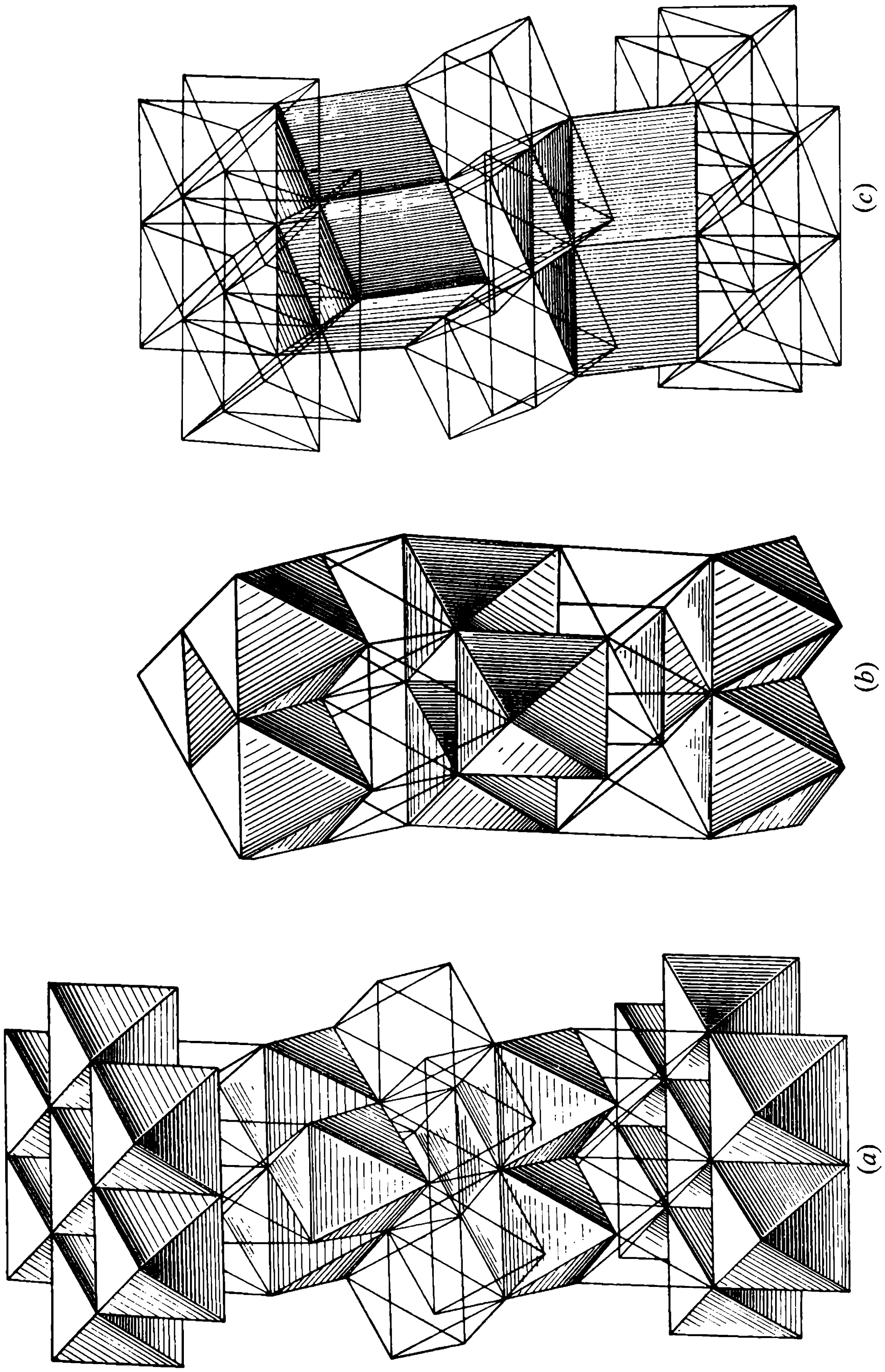
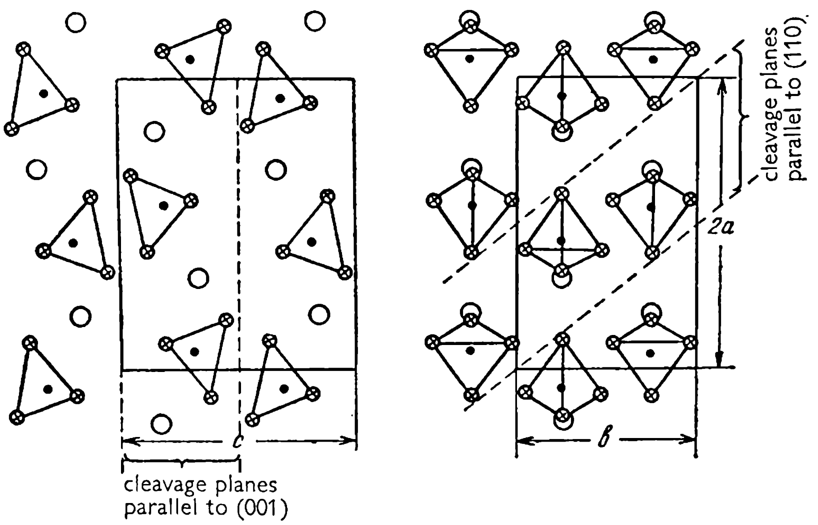


FIG. 270. Layer-type structures: (*a*) cadmium chloride; (*b*) cadmium iodide; (*c*) molybdenite.

and others have more complicated and varied shapes and various point symmetries. This also applies to the molecules H_2O, NH_3 etc. which can also be structural elements of a crystal. Crystals of complex compounds have structures of various types; molecular, layer or coordination, but in spite of the extra complexity there may still be a relation between cleavage and structure since the cleavage planes are such that they divide but do not cut through the structural elements. This is shown in Fig. 271 in which the traces of the cleavage planes are shown dotted on projections of the crystal structure of barite $BaSO_4$.


c
cleavage planes parallel to (001)
b
2a
cleavage planes parallel to (110)

(001) projection (001) projection

FIG. 271. Cleavage planes and the structure of the barite crystal.

In conclusion, consider crystals of the coordination type with simple structures. Metal crystals with f.c.c. structures Cu, Ag, Au, Pt, Al and others, as has already been said, have extensive plasticity and cleavage does not occur. Metals with b.c.c. structure—Cr, α-Fe, W and so on when deformed by impact at low temperatures have perfect cleavage on the cube plane $\{100\}$ this being in contradiction to Bravais' rule since the most densely packed and consequently most widely spaced planes are $\{110\}$.

The diamond crystal exhibits cleavage on the octahedral $\{111\}$ planes which agrees with structural ideas since these planes intersect the minimum number of bonds. In the sphalerite crystal α-ZnS, zinc atoms and sulphur atoms occupy the same sites as the carbon atoms in diamond, but the cleavage planes are the rhombic dodecahedral planes $\{110\}$ and not the octahedral planes as would be expected from Bravais' rule. The difference from diamond is explained by noting that in zinc sulphide the bonds between the atoms are partially ionic and it is thus necessary to take account of electrostatic interactions between them. In zinc sulphide, the $\{111\}$ planes which are filled with zinc atoms alternate with planes composed entirely of sulphur atoms. As a result of the partially ionic character of the bond these planes carry surplus negative and positive charges and attract each other with stronger forces than the $\{110\}$ planes which are nearer together but which each contain an equal number of zinc and sulphur atoms and are thus electrically neutral. Fig. 272 shows the projection of the sphalerite structure on the cube face. The traces of the cleavage planes are shown dividing electrically neutral planes of $\{110\}$ type. The rule of maximum interplanar spacing is really satisfied for sphalerite if one corrects for ionic interaction since from the countless number of planes which satisfy the condition of electrical neutrality, the planes of {110} type have the largest interplanar spacing.

In crystals of composition AB_2 the planes may be filled according to a sequence AB | BAB | BAB | BA.... In this case the cleavage planes pass between planes with the same letters. This is supported by the cleavage of fluorite CaF_2 on {111}, cuprite Cu_2O on {111} and rutile TiO_2 on {100} and {110}. Fig. 273 shows the projection of the structure of rutile and the traces of the cleavage planes which divide planes with the same structure. Ionic crystals of the CsCl type, just like crystals of NaCl type, exhibit cleavage on the cube faces {100} while, judged by the magnitude of the interatomic spacing and electrical neutrality, the cleavage planes should be {110}. This example shows the necessity of taking fuller account of the interaction between the particles of the crystal if one is to establish the connection between structure and cleavage. It also shows that simple criteria such as the Bravais rule or that for the total charge on a plane are insufficient.

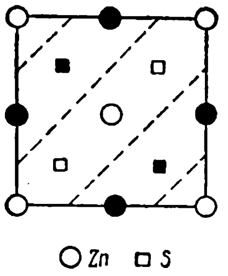


FIG. 272. Cleavage planes and the structure of sphalerite.

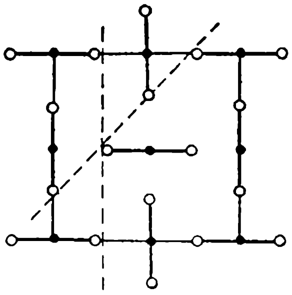


FIG. 273. Cleavage planes and the structure of rutile.

4. Plasticity of Crystals. Slip

Slip Systems

When crystals are subjected to mechanical forces above the yield stress, they acquire permanent deformations. Plastic deformation of crystals is in many cases, accompanied by the appearance of visible slip traces.

Investigation of the positions of slip lines and changes in the external shape of crystals shows that plastic deformation takes place by slip displacement on several crystallographic planes in definite crystallographic directions. The slip planes and slip directions in crystals are the planes and directions with least resistance to slip. The slip plane (hkl) and slip direction $[mnp]$ form a slip system in which the condition for parallelism gives $hm+kn+lp = 0$. The indexes of the slip planes and slip directions can be determined by X-ray and optical means.

In single crystals of close packed metals the slip planes are the planes with the densest packing of atoms and with the largest distance between them. The slip directions are the directions with the minimum interatomic spacing, that is the planes with the densest atomic packing. Table 74 gives the slip systems for various crystals and also values of τ_0 the critical shear stress.

In crystals of b.c.c. metals (iron, molybdenum) the slip directions are clearly defined as $\langle 111 \rangle$ and are the directions of densest atomic packing; there are no clearly

Table 74

Slip Systems and Critical Shear Stresses

Crystal	Lattice type	Slip system		τ_0, kg/mm^2
		$\{hkl\}$	$[mnp]$	
Cu	face-centred cubic	$\{111\}$	$[1\bar{1}0]$	0·100
Ag		$\{111\}$	$[1\bar{1}0]$	0·060
Au		$\{111\}$	$[1\bar{1}0]$	0·092
Ni		$\{111\}$	$[1\bar{1}0]$	0·580
Mg	hexagonal close-packed	(0001)	[100]	0·083
Zn		(0001)	[100]	0·094
Cd		(0001)	[100]	0·058
α-Fe	body-centred cubic	(110) (112) (123)	[111]	
Mo		(112) (110)	[111]	
β-Sn	tetragonal	(100) (110) (101) (121)	[001] [101]	0·16
NaCl	cubic	$\{110\}$	[101]	0·20
Bi	rhombohedral	(0001)	[100]	
Te	hexagonal	(0010)	[100]	

defined slip planes. The same can be said of the crystal of white tin, the structure of which still further departs from close packing of spheres. In the ionic crystal of sodium chloride the slip direction is ⟨110⟩ which is composed of ions of one sign. The slip planes are {110} which are among those with electrical neutrality. From amongst the crystals of simple substances with chain and layer type structures only tellurium and bismuth have significant plasticity. The slip elements for these crystals correspond to their structures. In the chain-type crystal of tellurium the slip planes are parallel to the axes of the chains; in the layer-type crystal of bismuth the slip plane is parallel to the plane of the layers. The slip directions coincide with the directions of the shortest spacings. In tellurium these directions are parallel to the axes of the chains.

The number of possible slip systems depends on the symmetry of the crystal. Thus in bismuth and tellurium there are three slip systems. In bismuth these systems are formed from one slip plane (0001) and three slip directions of ⟨100⟩ type, and in tellurium from three slip planes $\{10\bar{1}0\}$ and one slip direction ⟨100⟩ in each of them. Fig. 274*a* shows, on a stereographic projection, the positions of the slip elements in a hexagonal crystal with one slip plane I and three slip directions 1, 2 and 3 (as is the case for magnesium, zinc, cadmium and bismuth). Possible slip systems are I-1, I-2 and I-3. The stereographic projection in Fig. 274*b* shows the slip elements for a hexagonal or rhombohedral crystal with three slip planes of the type $\{10\bar{1}0\}$— I, II, III and three directions of slip 1, 2, and 3 (the case of tellurium). Possible slip systems are I-1, II-2, and III-3. Fig. 274*c* shows the stereographic projection of slip elements of a f.c.c. metal. The normals to the slip planes are I, II, III, IV (threefold axis of symmetry) the slip directions are 1, 2, . . . 6 (twofold axis of symmetry). The twelve possible slip systems in this crystal are as follows

$$\begin{matrix} \text{I},6; & \text{II},5; & \text{III},6; & \text{IV},5 \\ \text{I},1; & \text{II},2; & \text{III},3; & \text{IV},4 \\ \text{I},2; & \text{II},3; & \text{III},4; & \text{IV},1. \end{matrix} \qquad \ldots [16.19]$$

The active slip system is that for which the shear stress is a maximum. This depends on the type of stress and the orientation of the pole of the applied forces relative to the crystallographic axis.

Consider the simplest case of uniaxial extension or compression. Let the tensile force F form an angle ϕ with the normal to the slip plane N and an angle ψ with the slip direction T in this plane (Fig. 275). Then the shearing stress on the slip system (N, T) is

$$\tau = \frac{F}{S_0} \cos \phi \cos \psi. \qquad \ldots [16.20]$$

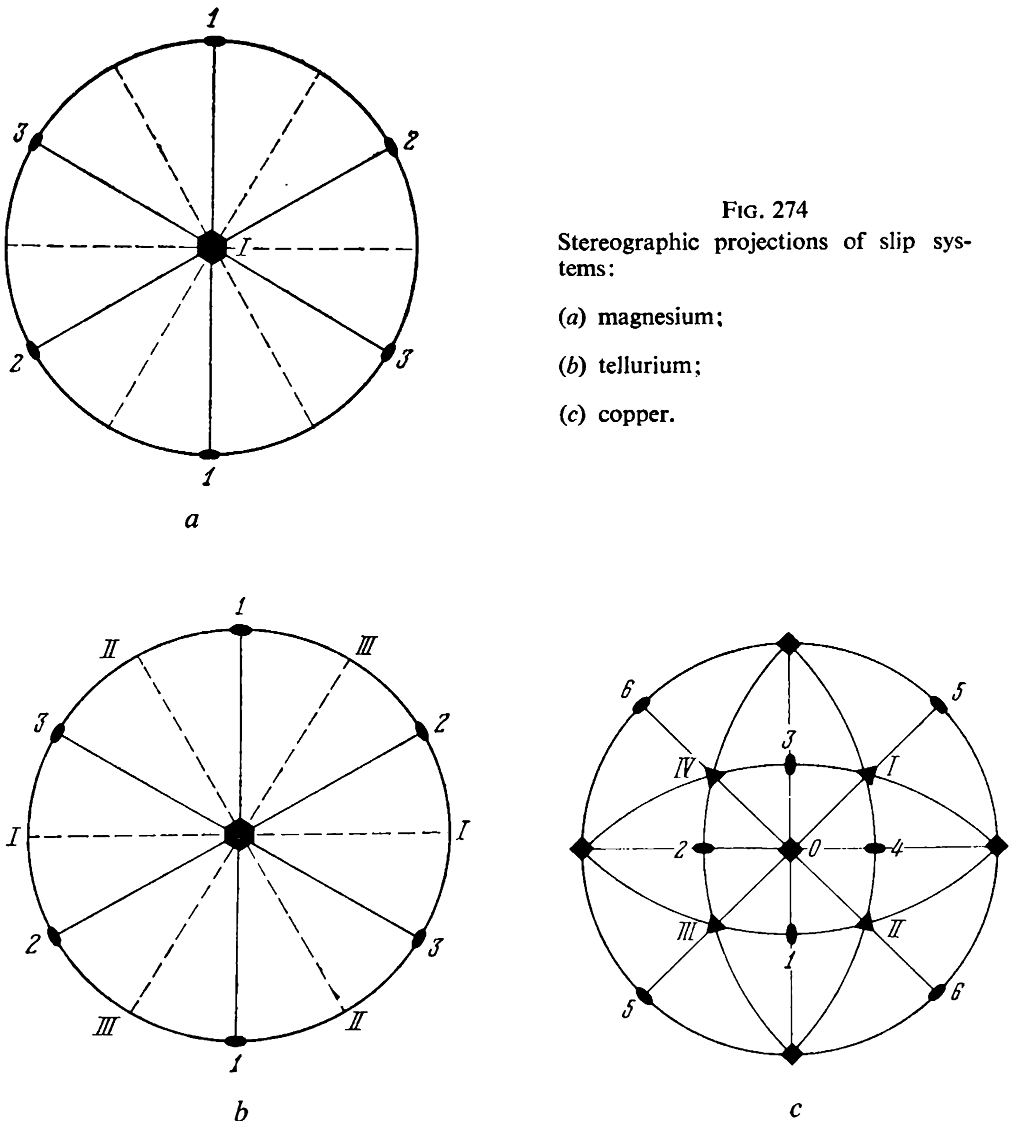


FIG. 274
Stereographic projections of slip systems:

(*a*) magnesium;

(*b*) tellurium;

(*c*) copper.

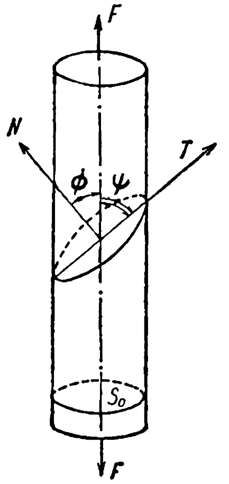


FIG. 275. Resolution of the tensile force into shear and normal components.

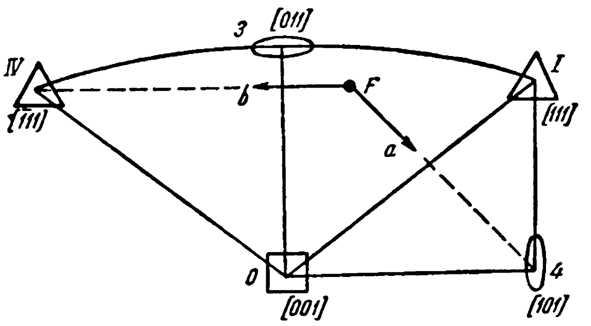


FIG. 276. Motion of the specimen axis under extension (arrow *a*) and compression (arrow *b*) of a crystal with f.c.c. structure.

The maximum value of the slip stress is that for the plane and direction which form an angle with the stress axis which is nearest to 45°. Formula [*20*] enables us to determine the critical shear stress for various systems of slip as a function of the point *F* (which marks the direction of the shearing force on the stereographic projection) and to determine the system with $\tau = \tau_{\max}$. On the projection we can find the operative system geometrically by drawing the line of intersection of the sphere with the cone of maximum stress (the cone with angle at 45° to an axis coincident with the direction of the force *F*). This line is a circle. Analysis of the possible variants leads to the following rule for finding the operative slip system in the f.c.c. lattice.

On Fig. 274*c* describe a spherical triangle in which lies the direction of the tensile stress. (This may be any of 48 equivalent triangles, for example OI3.) Now find two neighbouring triangles, the vertices of which are two- and three-fold axes of symmetry (Fig. 276), that is, the points 4 and IV which then define a system with maximum shear stress.

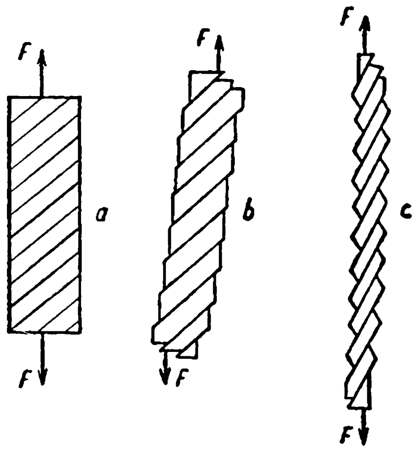


FIG. 277. Changes in the orientation of a crystal during extension.

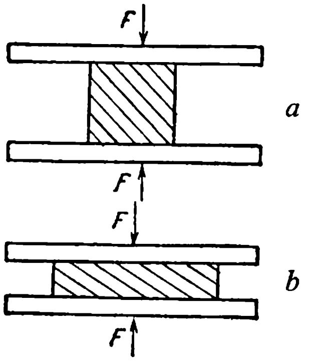


FIG. 278. Changes in the orientation of a crystal during compression.

Changes of Crystal Orientation during Deformation

Deformation of crystals is accompanied by changes of orientation. This is shown in Fig. 277*a* while Fig. 277*b* shows how pure slip causes the ends of the specimen to move further away from the stress axis. As a result, a couple is formed which produces local bending in zones adjacent to the clamped ends of the crystal. These bends produce changes in the orientation of the middle part of the crystal as a result of which the active slip directions move closer to the stress axis. In Fig. 276 this change of orientation of the crystal is shown by the little arrow *a*. In compression (Fig. 278*a*) the slip plane tends to set itself perpendicular to the compression axis (Fig. 278*b*) and in consequence the normal to the slip plane moves closer to the compression axis. Changes in the orientation of a crystal by compression are shown in Fig. 276 by the little arrow *b*. Thus, we can determine the slip system from changes of orientation of a crystal in two experiments (compression and extension).

Hardening Mechanisms

The force which has to be applied to a crystal to initiate slip is from [*20*] given by

$$F = \frac{\tau_0 S_0}{\cos\phi\cos\psi}. \qquad \ldots [16.21]$$

In otherwise similar conditions this force depends on the orientation of the crystal. Fig. 279 shows the dependence of the force F on the angle ϕ of inclination of the slip plane and on the product of the cosines in [*21*]. This curve is found to agree with experiment. The dependence of the tensile force F on the orientation of the crystal (the flow stress remaining unchanged) can be called *orientation hardening*. In fact, this phenomenon is not accompanied by any hardening of the crystal.

If we think of the critical shear strength τ_0 as analogous to the force of friction then we may expect it to increase with increase of external hydrostatic pressure P.

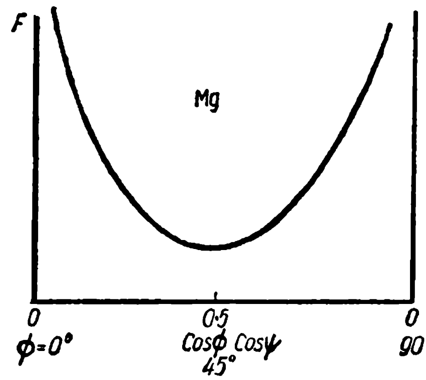


FIG. 279. Dependence of the critical shear stress on the crystal orientation.

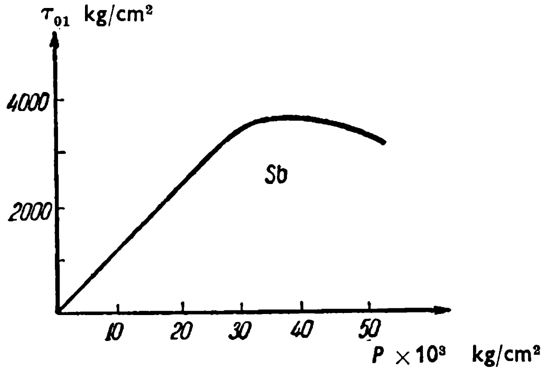


FIG. 280. Dependence of the critical shear stress for antimony on the external pressure.

Experiments on the extension of zinc and magnesium crystals at pressures up to 100 atm. gave no measurable effect. This result was interpreted as showing that there was an essential difference between resistance to slip in a crystal and a force of friction. Leaving aside the correctness of the above analogy we may note that in these experiments the pressure range was small compared with the internal pressure in condensed systems arising from the effect of cohesive forces. In Bridgman's experiments much higher pressures were used and showed that there was a significant effect of pressure on the shear stress. This can be called *pressure hardening*. Fig. 280 shows a curve obtained for antimony: notice the complex character of the curve $\tau_0 = \tau_0(p)$ which passes through a maximum (assuming that this maximum is not connected with peculiarities of the experiment). At 30,000 atm. paraffin has much the same viscosity as metals at normal temperatures, while graphite at 12,000 atm. begins to behave as an abrasive material. This unusual behaviour of graphite at high pressures is understandable if we remember that the C – C spacing in the atomic nets of graphite (1·41 Å) is less than the C – C spacing in diamond (1·54 Å). This is evidence that the strength of the bonds between the carbon atoms in graphite is greater than in diamond and agrees with the fact that at normal temperatures and pressures, graphite is a more stable modification of carbon than is diamond. The layer structure of graphite is the reason for its unique mechanical properties at ordinary pressures, but because of the high coefficient of compressibility in the direction perpendicular to the layers, those properties become less marked at high pressure.

The critical shear stress τ_0 depends on temperature. Fig. 281 shows this dependence for the hexagonal crystals magnesium and cadmium slipping on their basal planes. As the temperature is lowered the resistance to sheer increases, though not by much. For example the strength of cadmium increases by less than a factor of four when the temperature is lowered from a melting point down to 20° K. The plasticity of metallic crystals is retained right down to the lowest temperatures. This increase in the critical shear stress τ_0 of crystals as the temperature is lowered can be called *thermal hardening*.

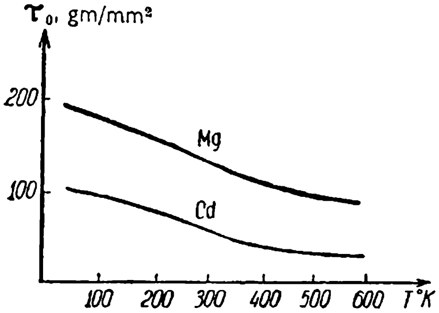


FIG. 281. Dependence of τ_0 on temperature for magnesium and cadmium.

A change in the chemical composition of a crystal produced by introducing impurity atoms into the crystal lattice, as in solid solutions, may lead to a large

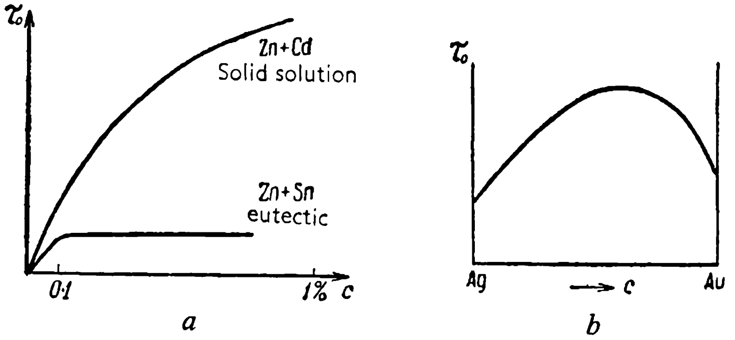


FIG. 282. Dependence of τ_0 on alloy concentration for a eutectic and two solid solutions.

increase in the critical shear stress. This can be called *alloy hardening*. Fig. 282*a* shows the change of τ_0 in crystals of zinc as a function of cadmium and tin content. When the concentration of cadmium in solid solution is progressively increased, the critical shear stress rises continuously. The hardening effect of tin atoms is observed only in the region of small concentration, since the equilibrium diagram of the zinc-tin system is of the eutectic type with small solubility. When the tin content is increased above the solubility limit, the tin atoms no longer go into solid solution and thus have no effect on the properties of the solid solution. Fig. 282*b* shows the hardening produced by alloying in the case where a continuous range of solid solutions are formed (the silver-gold alloys). The maximum hardening occurs for alloys with 50% of each of the components.

Plastic deformation destroys the regularity of the crystal structure and leads to large increases in the critical shear stress. Fig. 283*a* shows curves of the dependence of the critical shear stress of aluminium on the amount of deformation at various temperatures. The initial value of τ_0 is determined by the points where the curves intersect the ordinate axes. The curves showing the hardening effect of mechanical deformation on aluminium can be described approximately by parabolic equations $\tau = \tau_0 + c\sqrt{\lambda}$. Fig. 283*b* shows the hardening curves for zinc at various temperatures. These curves, unlike those for aluminium, are better expressed by a linear law. Hardening resulting from plastic deformation is called *work hardening*.

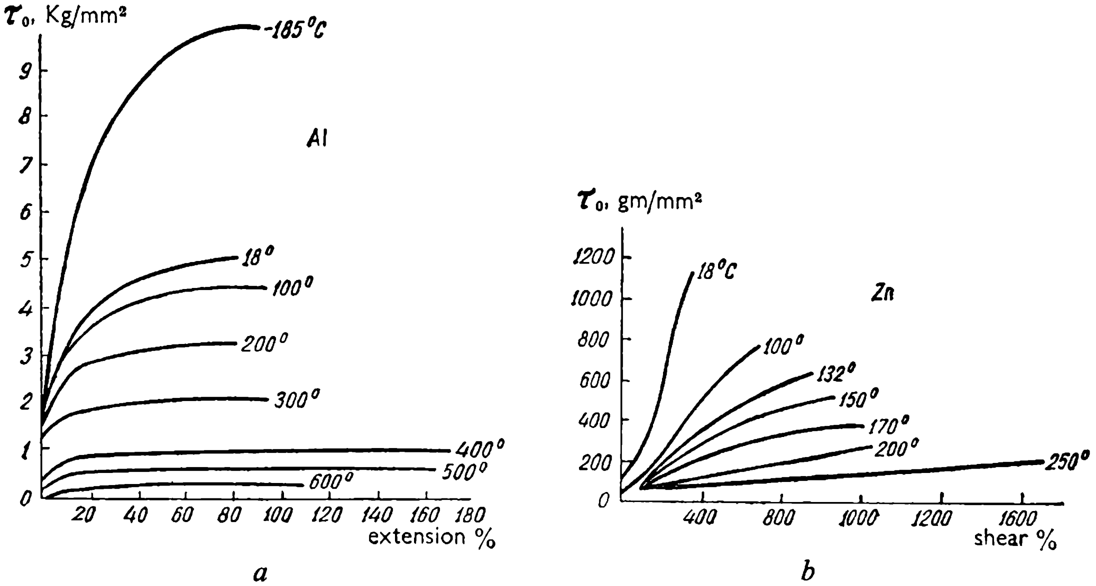


FIG. 283. Dependence of τ_0 on temperature for aluminium and zinc.

Hot and Cold Working

The dependence of the hardening curves on temperature and the reduction of the hardening effect at higher temperatures, show that in plastic deformation of crystals there are two types of process taking place and that these act in opposite directions. Deformation of crystals leads to the accumulation of residual strains and internal stresses which increase with the amount of deformation. An increase of temperature is accompanied by an increase in thermal energy of the atoms and an increase of their mobility and this leads to some reduction of the internal stresses and strains in the lattice. The influence of temperature in neutralising the effects of plastic deformation increases as the temperature is raised. At sufficiently high temperatures this softening which is produced by thermal agitation may almost completely cancel out the effects of work hardening. We may thus distinguish two temperature regions: the low temperature region where the effect of thermal agitation is small, and the mechanical hardening of the crystal, which is connected with a metastable state of the lattice, may be retained almost indefinitely. This region is thus described as *cold working*. The high temperature region, where thermal agitation leads to complete removal of the effects of mechanical working is called the region of *hot working*. Hot and cold working of metals and alloys is extensively used in metallurgical processes.

The Energy of the Internal Stresses

In plastic deformation of metals, the internal energy increases by an amount

$$\Delta U = \delta A - \delta Q. \qquad \ldots [16.22]$$

Here δA is the work done by the external forces to produce the deformation and can be measured from the stress-strain curve. The second term δQ is the heat produced during plastic deformation and can be measured calorimetrically.

Fig. 284 shows the change of internal energy of cold worked copper as a function of the amount of deformation. The curve shown is related to the hardening curve (Fig. 283*a*): at small deformations when the rate of hardening is large, the rate of change of internal energy is also large. At large strains when the hardening curve becomes saturated the energy of the residual strains also tends to some limit. Values of the energy ΔU are given below for various metals: ΔU is about 30-70 calories per gm.-atom.

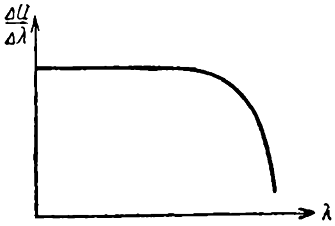


FIG. 284. Dependence of the rate of change of internal energy on the amount of cold work.

Energy of Residual Stresses in Some Plastically Deformed Metals

Metal	Al	Fe	Ni	Cu	Brass
ΔU cal/gm	1·1	1·2	0·78	0·50	0·49

Under adiabatic conditions, the energy of the internal stresses is sufficient to heat the body by several degrees.

Classification of Internal Stresses

Davidenkov's classification of plastic deformation of crystals distinguishes three kinds of crystal strain, to which correspond internal stresses of types I, II, and III.

Strains of first and second kind are small displacements of the atoms of the crystal relative to each other. Within the unit cell the corresponding atomic displacements are only a small fraction of the smallest interatomic spacing (of the order of 1/1000 and certainly not greater than 1/100). These displacements can take place without going far from the local potential minima which correspond to the equilibrium positions of the atoms, that is without transition through the potential barriers which separate nearest neighbours. Strains of types I and II are accompanied by internal stresses of elastic character. Strains of type III arise when atoms make the transitions through the potential barriers, that is when atoms move by distances which are not less than half the interatomic spacing and may be much larger. At large displacements, the atoms of neighbouring layers may not fall completely into the minima which correspond to the equilibrium lattice. At the end of a slip process, a fraction of the atoms occupy intermediate positions of equilibrium (one might say that 'they remain in the passes') which may be situated at distances from the equilibrium positions which are a large fraction of the interatomic spacings. Strains of type III arise as a result of passing through the elastic (or flow) limit. They are metastable defects in the crystal structure and are retained when the external forces are removed. Thus they are responsible for the irreversibility of plastic deformation. Stresses of type III correspond to these local strains and can be called superelastic stresses.

The above mentioned stresses are distinguished by the size of the volumes in which they are operative. Strains of type I extend throughout the whole volume of the specimen or metal article and are uniformly distributed over the grains. Examples are: elastic bending of a beam or a rail and deformation of a piece of quenched metal. In individual parts which are small compared to the total volume of the article, but which contain a large number of micro-crystals, deformation is homogeneous and is accompanied by a change of the interplanar spacing d by Δd. Strains of type I are shown up by X-rays as a movement of the lines produced by a change of the angle of reflection $\Delta\theta$.

Strains of type II are elastic strains non-uniformly distributed among the microcrystals, or amongst definite parts of the microcrystals of a polycrystalline specimen. These strains arise as a result of plastic deformation of a polycrystal (rolling, wire-drawing, forging and so on). The non-uniformity of deformation is a result of the crystallographic character of plastic deformation of individual crystals and of their different relative orientations in the polycrystal. Strains of type II extend over regions which contain a large number of unit cells. They can be detected by broadening of X-ray lines since the diffraction spots from the individual uniformly deformed parts of the microcrystals are displaced by angles $\pm\Delta\theta_i$, which can be of either sign corresponding to extended and compressed parts and changes of interplanar spacing $\pm\Delta d_i$. Strains of type I can be thought of as oriented strains of type II which extend over larger groups of microcrystals.

Strains of type III extend over a volume of several unit cells, and to them there correspond atomic displacements which destroy the regularity of the lattice. Strains of type III are shown up as a weakening of the X-ray line intensity by amounts which are similar to those due to thermal vibrations. The difference is that lattice disorder produced by thermal vibrations has a dynamic character and the instantaneous displacements of atoms from their positions of equilibrium change continuously with time, with the centres of vibration of the atoms forming a regular lattice in space. Displacements of the atoms connected with stresses of type III have a static character and to some degree correspond to the state of a lattice as produced by thermal agitation but frozen at some moment of time. Centres of thermal vibration of a deformed lattice are fixed in space but do not form a completely regular lattice.

In Fig. 285 lattice strains of yet another kind are shown schematically. When the angle of bending is large, elastic bending changes into plastic and this is accompanied by slip which results in the crystal being broken up into a number of elastically bent packets with a corresponding reduction in the total elastic energy of bending. In each elastically bent packet there are stresses of type II, while on the boundaries between the packets where the largest departures from regularity of the structures are concentrated there are stresses of type III. The number of atoms in neighbouring packets per unit length along the boundary is different, being arranged like the graduations on a vernier.

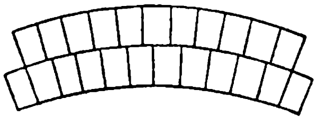


FIG. 285. Model of lattice strain produced by plastic bending.

We can estimate the magnitude of internal stresses from X-ray data on deformed lattices. The energy of elastic stresses is determined by the formula

$$\Delta U = \tfrac{1}{2}E\lambda^2 = \tfrac{1}{2}E\left(\frac{\Delta d}{d}\right)^2, \qquad \ldots [16.23]$$

where Δd is obtained from the line displacement for stresses of type I and by the line broadening for stresses of type II. Calculation, using formula [*23*], gives a value $\Delta U = 0{\cdot}015$ cal/gm for stresses of type II in copper. This is only a few percent of the energy stored in a deformed metal. The energy of type III stresses can be estimated by a formula which is similar to that for the heat which goes to raise the temperature of a body by ΔT. This is

$$\Delta U' = \tfrac{1}{2}c\Delta T, \qquad \ldots [16.24]$$

where c is the specific heat, and the coefficient $\frac{1}{2}$ arises since in calculating the energy of static stresses we must take account only of potential energy. The effective temperature ΔT is determined by comparing the weakening of the intensity of a line as a result of plastic deformation with the weakening produced by raising the temperature. Calculation using formula [*24*] gives a value for $\Delta U'$ which agrees in order of magnitude with the energy of the internal stresses measured calorimetrically. This shows that internal stresses of type II contribute only about 2% of the total stored energy while stresses of type III contribute the remaining 98%.

The File Model

Consider the dynamics of slip in a crystal. Assume that the crystal is ideal and that slip takes place simultaneously on all slip planes (the file model). In this case the initial critical shear stress, as given by Frenkel is

$$\tau_0 = G\lambda_0, \qquad \ldots [16.25]$$

where G is the shear modulus and λ_0 is the elastic deformation at the onset of slip. Take $\lambda_0 = 0{\cdot}1$ of the interatomic spacing and $G = 3{\cdot}5 \times 10^5$ kg./cm^2 (for zinc) and we obtain $\tau_0 = 350$ kg/mm^2, which is several orders of magnitude greater than the experimental value. Thermal agitation lowers the critical shear stress but even at absolute zero, the critical shear stress is about a thousand times less than the value obtained from the ideal crystal model.

Dislocations

Suppose that in a crystal there are defects of vernier type (Taylor, 1934), such that the number of atoms in one part of the crystal is greater than the number of atoms opposite them in another part. Fig. 286*a* shows the arrangement of atoms in an ideal lattice and the periodic variation of potential energy along a given conventional row, for example, *BB*

$$U = 2A \cos 2\pi \frac{x}{a}. \qquad \ldots [16.26]$$

Fig. 286*b* shows the arrangements of atoms in the damaged region, where it is assumed that there are similar arrangements of atoms on all parallel lattice planes

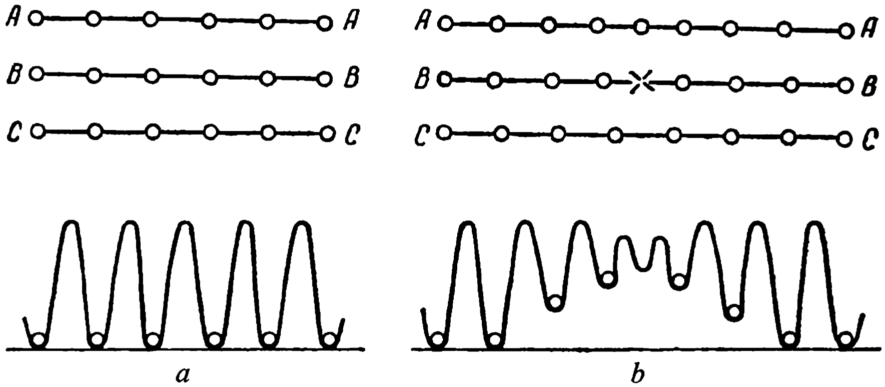


FIG. 286. Potential energy curve: (*a*) in an ideal crystal; (*b*) in the region of a dislocation.

and that there is a defect in the crystal produced by the introduction of an additional half plane. The intersection of the edge of the half plane with the plane of the drawing is indicated by a cross. This type of defect is called an *edge dislocation*. The potential energy curve for the parts which contain respectively N and $N+1$ atoms is of the beat frequency type

$$U' = A \cos 2\pi \frac{x}{a}\left(\frac{N+1}{N+\frac{1}{2}}\right) + A \cos 2\pi \frac{x}{a}\left(\frac{N}{N+\frac{1}{2}}\right). \qquad \ldots [16{\cdot}27]$$

In the centre of the dislocation there is a shallow unoccupied minimum separated from the neighbouring minima by small barriers. The shape of the potential curve shows that the dislocation can diffuse by thermal fluctuations along the slip plane. When shearing forces are applied to the crystal, the curve of potential energy, as is shown in Fig. 287, becomes asymmetrical

$$U'' = A \cos 2\pi \frac{x-\delta}{a}\left(\frac{N+1}{N+\frac{1}{2}}\right) + A \cos 2\pi \frac{x}{a}\left(\frac{N}{N+\frac{1}{2}}\right), \qquad \ldots [16.28]$$

as a result of which the dislocation undergoes a directed displacement corresponding to plastic slip of one part of the crystal relative to the other. The magnitude of the local minima in the centre of the dislocation is much less than in an ideal lattice and this corresponds to a lowering of the critical shear stress necessary to move the dislocation in the slip plane. The difference from an ideal crystal is that in the defect crystal which includes dislocations, slip proceeds not by the simultaneous movement of all the atoms on the slip plane but by small groups of atoms moving in turn. This corresponds to motion of a dislocation within the crystal lattice.

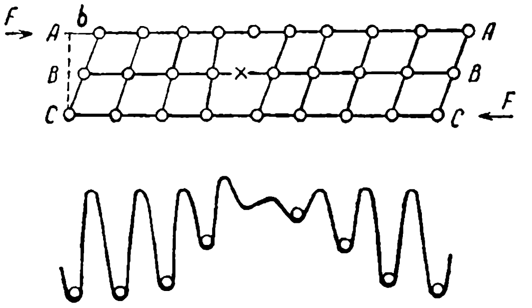


FIG. 287
Potential energy curve in the dislocation zone under the influence of a shearing stress.

Calculation of the energy of an edge dislocation gives a value of the order of 10^{-5} erg/cm length. The additional energy in 1 cm^3 of a deformed metal $\Delta U \approx 10$ $\text{cal/cm}^3 = 5 \times 10^7$ erg/cm^3. Hence 1 cm^3 of deformed metal contains $\sim 5 \times 10^{12}$ cm of dislocation line or on its faces $\sim 2 \times 10^{12}$ dislocations emerge per cm^2. Taking the atomic diameter as 3 Å there are 10^{15} atoms per cm^2 and consequently 10^3 atoms to each dislocation. Each dislocation is connected with the displacement of a large number of atoms, so the total number of displaced atoms in a plastically deformed metal will be large and this agrees with X-ray results which show that at large deformations one can on the average consider every eighth atom in the lattice to be displaced. The excess energy of a deformed metal can be identified with the energy of the dislocations.

We may distinguish two types of edge dislocations and call them positive and negative (Fig. 288). Positive dislocations are those such that the upper half of the crystal is in tension and the lower part in compression. In negative dislocations the positions of the tensile and the compressive parts are reversed. It is not difficult to see that dislocations within a crystal must interact with each other: when dislocations of the same sign come close together the displacement energy of the atoms increases since parts with deformations of the same sign (tension or compression) overlap each other. Thus dislocations of the same sign repel each other and conversely dislocations of opposite sign attract each other since, when they become closer together their energy of interaction is reduced. When dislocations of opposite sign merge, stresses

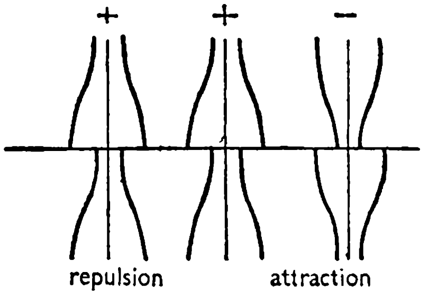


FIG. 288. Positive and negative dislocations.

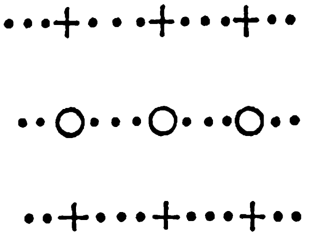


FIG. 289. Ordered arrangement of dislocations.

of kind III are relieved and the lattice is healed. Quantitative calculations of the dependence of the energy of dislocations on the distance between them shows that dislocations interact with each other with a force which is reciprocally proportional to the distance between them $1/r$. In equilibrium and in the absence of external forces dislocations tend to form an ordered array within a crystal (Fig. 289).

The resistance of the crystal to shear is connected with repulsion of dislocations of the same sign in the same slip planes and the attraction of dislocations of opposite sign in different slip planes. Work hardening of the crystal is explained by increase of the density of dislocations and increase of the forces of interaction between them. Consider now a crystal with one slip system. Let there be n_1 active slip planes per unit length along the normal and on each of these planes let there be n_2 dislocations. The dislocation density in this case is $N = n_1 n_2$. When dislocations move along the slip planes this produces displacement of individual slip packets. Let L be the mean free path of a dislocation in a crystal. When all the dislocations on a given plane move, the slip packet moves by $\Delta l = an_2L$, where a is the interatomic spacing. Since the width of the packet is $1/n_1$ the total shear in the crystal is

$$\lambda = \Delta l n_1 = an_1n_2L = aNL. \qquad \ldots [16.29]$$

If the number of effective slip planes in the deformation process remains constant, then n_1 = const. and the mean spacing between dislocations on the slip planes is

$$r \approx \frac{1}{n_2} = \frac{\text{const}}{N}. \qquad \ldots [16.30]$$

The resistance to shear in a crystal is determined by the force of interaction between the dislocations which is reciprocally proportional to the mean spacing between them

$$\tau \approx f = \frac{c}{r} = cN = c'\frac{\lambda}{L}. \qquad \ldots [16.31]$$

In this case the shearing stress is proportional to the shear. When the number of effective slip planes is increased, slip is proportional to the total number of dislocations, that is, for a uniform distribution of dislocations throughout the cross-section of the crystal $n_1 = n_2 = \sqrt{N}$.

In this case the mean spacing between dislocations is

$$r \approx \frac{\text{const}}{\sqrt{N}}, \qquad \ldots [16.32]$$

and instead of [*31*] we obtain

$$\tau = \frac{c}{r} = c\sqrt{N} = c'\sqrt{\frac{\lambda}{L}}, \qquad \ldots [16.33]$$

which gives a parabolic dependence of the shear stress on the shear strain.

The difference between the laws of hardening for hexagonal and cubic close-packed metals is connected with differences in the mechanism of deformation. One such difference is that when single crystals of hexagonal metals are extended the pattern of visible traces of slip planes is developed almost as soon as the extension begins, whereas in the face-centred cubic metals new traces of slip planes appear continuously as the extension increases. The reason for this difference is that in the hexagonal metals there is one slip system while in the f.c.c. metals there is usually double slip, even in the early stages.

Comparison of the theoretical hardening curves with experiment enables us to show that L—the mean free path of dislocations—is about 10^{-4} cm. This is the same as the mean dimensions of mosaic blocks in metals. Thus we may suppose that the boundaries of mosaic blocks are impassable obstacles for dislocations and act as collectors for them.

An important problem in the theory of plastic deformation of crystals is the origin of dislocations. Now the energy associated with dislocations greatly exceeds the mean energy per atom of thermal agitation, so dislocations cannot arise from thermal fluctuations. Alternatively if dislocations were easily formed by thermal agitation they would just as easily be destroyed by it, so there would be no mechanical hardening. We must thus assume that there is a certain number of dislocations within a crystal when deformation starts. However when plastic deformation takes place, the movement of dislocations and their departure into the mosaic block walls must continuously decrease the number of dislocations. This would quickly lead to exhaustion of the supply of dislocations and to a rate of hardening much larger than is observed by experiment. Dislocations must thus arise during the process of plastic deformation, the energy being supplied by the external forces. The mechanism of formation of dislocations depends on non-uniform loading conditions even in the case of an ideal crystal. If a crystal were absolutely hard then a stressed state could be instantaneously distributed throughout the entire crystal giving rise to a state of uniform stress. Now, stresses are distributed throughout a crystal by elastic waves with finite velocity. Inevitable non-uniformities in applying pressure to the surface of the crystal and also inevitable non-uniformities and defects in the crystal such as cracks, impurities, crystal boundaries, density fluctuations etc., which have existed in the crystal since it was formed, all lead to the formation of defects of the type of edge and screw dislocations.

Recently it has been possible to give direct evidence of dislocations in a way which supports the correctness of the model (Fig. 286*b*) put forward 25 years previously. Fig. 290*a* shows an electron microscope picture from a thin crystalline plate of platinum phthalocyanide on which the traces of planes ($d = 12$ Å) of platinum atoms can be seen. This picture also shows an edge dislocation which appears as the sudden end of the trace of one of the atomic planes. The corresponding region is shown schematically in Fig. 290*b*. Still greater resolving power can be obtained by forming Moiré figures in the electron microscope. These are formed when two thin crystalline platelets with slightly different lattice spacings are superposed. Fig. 290*c* shows a Moiré electron micrograph from superposed Au and Pd films. The dislocations shown are similar to those revealed in platinum pthalocyanide by direct electron microscopy.

5. Mechanical Twinning

External forces applied to crystals may produce an irreversible deformation different from slip. This type of deformation arises from the appearance of macroscopic uniform regions of the crystal with orientation very different from the orientation of the original lattice. Fig. 291 shows twinned regions formed when single metal

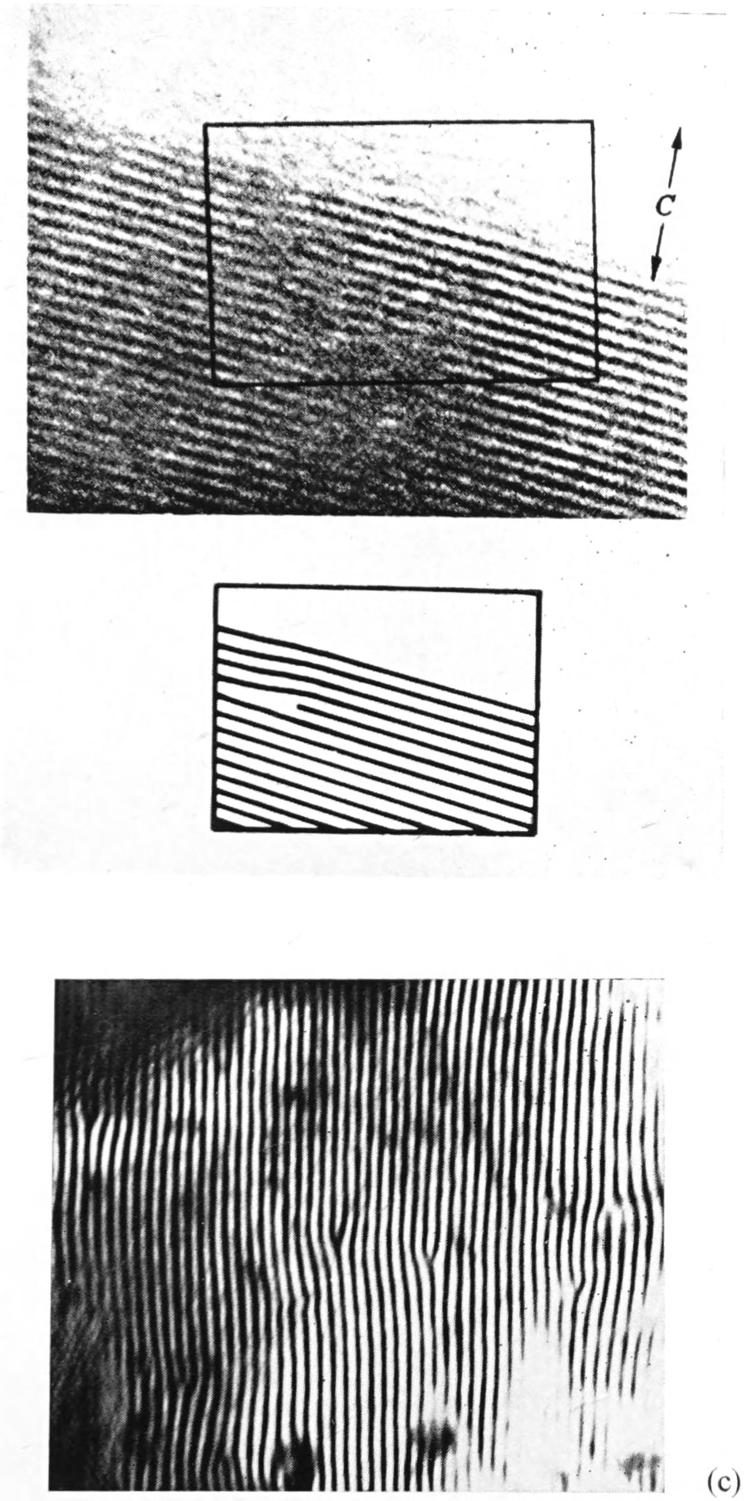


FIG. 290. Electronmicrograph of a dislocation.

(a), in pththalocyanine: the layer spacing is about 12Å. (J. W. Menter).

(b), Schematic drawing of the dislocation shown in (a).

(c), Electronmicrograph of a dislocation shown by a Moiré pattern of overlapping (111) Au and Pd films. The Moiré spacing is about 35Å and corresponds to an effective magnification of about 400,000. (D. W. Pashley, J. W. Menter and G. A. Bassett).

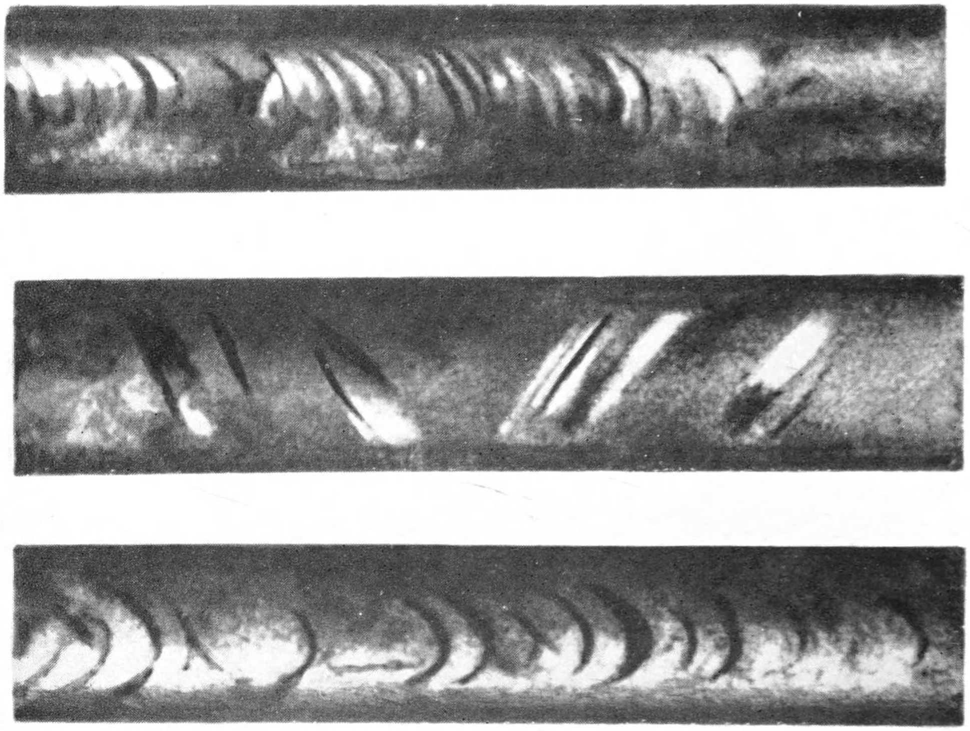


FIG. 291. Mechanical twins, formed when single metal crystals are extended.

crystals are extended. Whether extension of a crystal leads to slip or twinning depends on the conditions of the experiment. If the slip plane is unfavourably oriented relative to the tensile stress then increase in the external load leads to a large increase in the elastic stored energy. As a result of orientation hardening, this energy cannot be got rid of by means of plastic slip and in some crystals the energy is released by rotating the lattice into a new position and the formation of a mechanical twin. As experiment shows, the twinned part of a crystal is a mirror image of the lattice of the original crystal in some plane called the twinning plane. The twinning plane is a crystallographic plane usually with small indices.

The mechanism of twinning is different from the mechanism of slip. Consider this, taking as example twinning of a simple translational lattice. Fig. 292*a* shows the movements of atoms in simple shear (slip) along the slip direction *t*. The displacements of the slipping points relative to each other in this case are integral multiples of the lattice periods na_t along the slip directions. Here *n* is a whole number measuring the slip of neighbouring points and taking the required value. Fig. 292*b* shows atom movements in shear twinning taking place along the twinning direction η. In this case neighbouring atomic planes are displaced relative to each other by a strictly determined amount ma_η, which is less than the lattice period along the twinning direction a_η (m = constant<1).

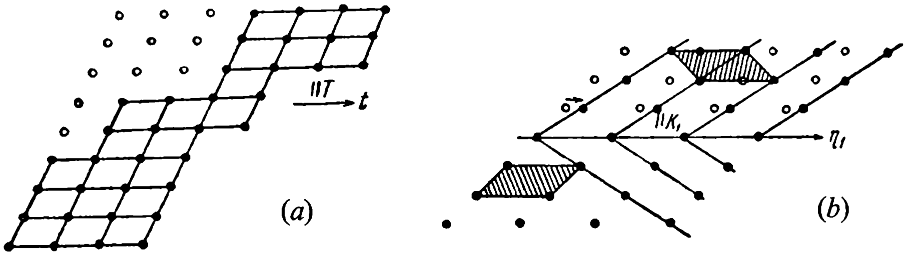


FIG. 292. Atomic movements: (*a*) in slip; (*b*) in twinning.

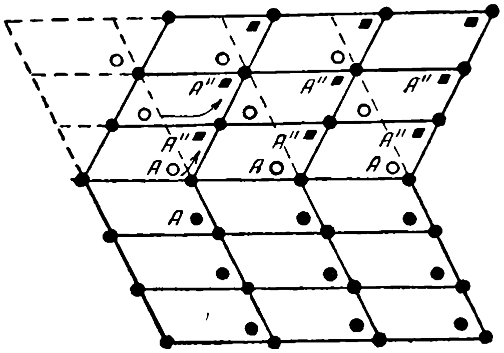


FIG. 293. The structure of a twin in calcite.

In crystals with a layer-type structure containing several particles in the unit cell ($CaCO_3$ is such a crystal) the formation of a twin, except by parallel motions, requires it to be possible for structural elements to rotate. Consider, as was first done by Gogoberidze, the simplest molecular crystal (Fig. 293). As well as displacements of the centres of the molecules along the twinning direction there is also rotation of the molecules round axes lying in the twinning plane perpendicular to the twinning direction.

In hexagonal metals such as zinc there is one slip plane (0001) and six twinning planes $\{10\bar{1}2\}$. If ordinary slip is impossible or prevented, then twinning takes place and this is accompanied by the formation of minima on the stress-strain curve corresponding to the relief of stresses. When a twin is formed plastic deformation takes place by slip.

APPENDIX 1

PROPERTIES OF THE ELEMENTS

Table I

First ionization potential of atoms, eV

H 13·56																	He 24·48
Li 5·37											Be 9·30	B 8·28	C 11·24	N 14·51	O 13·57	F 17·46	Ne 21·47
Na 5·09											Mg 7·63	Al 5·94	Si 8·14	P 10·43	S 10·42	Cl 13·01	Ar 15·68
K 4·32	Ca 6·25	Sc 6·7	Ti 6·81	V 6·74	Cr 6·7	Mn 7·41	Fe 7·83	Co 7·8	Ni 7·6	Cu 7·67	Zn 9·37	Ga 5·97	Ge 8·10	As 10·05	Se 9·75	Br 11·82	Kr 13·94
Rb 4·18	Sr 5·68	Y 6·6	Zr 6·92	Nb	Mo 7·2	Tc	Ru 7·5	Rh 7·7	Pd 8·30	Ag 7·58	Cd 8·94	In 5·76	Sn 7·54	Sb 8·35	Te 8·89	I 10·43	Xe 12·08
Cs 3·86	Ba 5·21	La 5·59	Hf	Ta	W	Re 7·8	Os	Ir	Pt 8·8	Au 9·20	Hg 10·41	Tl 6·08	Pb 7·37	Bi 7·25	Po	At	Rn 10·69
Fr	Ra 5·21	Ac															
				Ce 6·54	Pr 5·76	Nd 6·31	Pm	Sm 6·55	Eu 5·64	Gd 6·65	Tb 6·74	Dy 6·82	Ho	Er	Tm	Yb 7·06	Lu
				Th	Pa	U	Np	Pu	Am	Cm							

Table II
Valencies

H 1																	He
Li 1											Be 2	B 3	C 2, 4	N 3	O 2	F 1	Ne
Na 1											Mg 2	Al 3	Si 4	P 3, 5	S 2, 4, 6	Cl 1, 3, 5	Ar
K 1	Ca 2	Sc 3	Ti 2, 3, 4	V 5	Cr 3, 6	Mn 2,3,4,6,7	Fe 2, 3	Co 2, 3	Ni 2	Cu 1, 2	Zn 2	Ga 3, 1, 2	Ge 4, 2	As 3, 5	Se 2, 4, 6	Br 1	Kr
Rb 1	Sr 2	Y 3	Zr 4, 2, 3	Nb 5	Mo 6	Tc	Ru 6, 4, 8	Rh 3, 4	Pd 2, 4	Ag 1	Cd 2	In 3, 1, 2	Sn 4, 2	Sb 3, 5	Te 2, 4, 6	I 1	Xe
Cs 1	Ba 2	La 3	Hf 4	Ta 5	W 6	Re 6, 7	Os 6, 4, 8	Ir 3, 4	Pt 2, 4	Au 1, 3	Hg 2, 1	Tl 1, 3	Pb 2, 4	Bi 3, 5	Po	At	Rn
Fr	Ra	Ac															
				Ce 3, 4	Pr 3, 4	Nd 3	Pm	Sm 2, 3	Eu 2, 3	Gd 3	Tb 3, 4	Dy 3	Ho 3	Er 3	Tm 3	Yb 2, 3	Lu 3
				Th 4	Pa	U 4, 5, 6	Np	Pu	Am	Cm							

Table III

Density of the Solid Phase, gm/cm^3

H 0·07																	He
Li 0·53											Be 1·84	B 1·7	C 2·25	N	O	F	Ne
Na 0·97											Mg 1·74	Al 2·69	Si 2·2	P 1·8	S 2·0	Cl	Ar
K 0·86	Ca 1·55	Sc	Ti 4·5	V 5·7	Cr 7·1	Mn 7·3	Fe 7·86	Co 8·8	Ni 8·8	Cu 8·93	Zn 7·14	Ga 5·9	Ge 5·40	As 5·72	Se 4·80	Br	Kr
Rb 1·52	Sr 2·60	Y 4·57	Zr 6·53	Nb 8·50	Mo 10·2	Tc —	Ru 12·28	Ph 12·3	Pd 11·5	Ag 10·5	Cd 8·64	In 7·25	Sn 7·28	Sb 8·69	Te 6·25	I	Xe
Cs 1·87	Ba 3·8	La 6·15	Hf 13·3	Ta 16·2	W 19·1	Re 21·2	Os 22·48	Ir 22·4	Pt 21·4	Au 19·3	Hg 14·2	Tl 11·85	Pb 11·34	Bi 9·80	Po	At	Rn
Fr	Ra	Ac															
				Ce 6·80	Pr 6·47	Nd 7·0	Pm	Sm 7·5	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
				Th 11·5	Pa —	U 19·6	Np	Pu	Am	Cm							

Table IV
Heats of Sublimation, kcal/mole

H																	He
Li 39·0											Be 75	B	C	N 1·86	O	F	Ne 59
Na 25·9											Mg 36·3	Al 55	Si	P	S	Cl	Ar 2·03
K 19·8	Ca 47·8	Sc 70	Ti 100	V 85	Cr 88	Mn 74	Fe 94·0	Co 85·0	Ni 85·0	Cu 81·2	Zn 27·4	Ga 52	Ge 85	As 30·3	Se	Br	Kr
Rb 18·9	Sr 47	Y 90	Zr 110	Nb >68	Mo 160	Tc	Ru 120	Rh 115	Pd 110	Ag 68	Cd 26·8	In 52	Sn 78	Sb 40	Te	I	Xe
Cs 18·8	Ba 49	La 90	Hf >72	Ta >97	W 210	Re	Os 125	Ir 120	Pt 127	Au 92	Hg 14·6	Tl 40	Pb 47·5	Bi 47·8	Po	At	Rn
Fr	Ra 72·7	Ac															
				Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
				Th 177	Pa	U 220	Np	Pu	Am	Cm							

Table V

Melting Points, °C

																H −259	He
Li 180											Be 1278	B 2300	C 3900	N −210	O −219	F −223	Ne
Na 99·7											Mg 650	Al 660	Si 1414	P 44	S 113	Cl −103	Ar
K 63·6	Ca 850	Sc	Ti 1800	V 1715	Cr 2000	Mn 1250	Fe 1530	Co 1490	Ni 1452	Cu 1083	Zn 419	Ga 30·2	Ge 958	As 817	Se 217	Br −7·3	Kr
Rb 39·0	Sr ~800	Y	Zr 1900	Nb 1950	Mo 2600	Tc	Ru 2000	Rh 1970	Pd 1553	Ag 960	Cd 320	Zn 155	Sn 232	Sb 630	Te 453	I 114	Xe
Cs 28·5	Ba 850	La 826	Hf 2230	Ta 3027	W 3390	Re 3170	Os 2500	Ir 2360	Pt 1771	Au 1063	Hg −39	Tl 303	Pb 327	Bi 271	Po	At	Rn
Fr	Ra	Ac															

Ce 780	Pr 950	Nd 850	Pm	Sm 1350	Eu 1150	Gd	Tb	Dy	Ho	Er	Tm	Yb ~1800	Lu

Th 1850	Pa	U 1129	Np	Pu	Am	Cm

Table VI
Electrical Resistivities at 0° C, ohm.cm × 10^6

H																	He
Li 8·5											Be 5·5	B	C	N	O	F	Ne
Na 4·3											Mg 3·2	Al 2·53	Si	P	S	Cl	A
K 6·1	Ca 4·0	Sc	Ti 8·2	V	Cr 2·6	Mn	Fe 8·9	Co 6·2	Ni 6·6	Cu 1·56	Zn 5·5	Ga 40·8	Ge	As 35·2	Se	Br	Kr
Rb 11·6	Sr 30·3	Y	Zr 41	Nb	Mo 4·38	Tc	Ru 11·6	Rh 4·65	Pd 10·2	Ag 1·47	Cd 6·7	In 8·35	Sn 10·4	Sb 35·3	Te	I	Xe
Cs 18·2	Ba 57·5	La 57·6	Hf 29·6	Ta 13·8	W 4·89	Re	Os 8·9	Ir 4·85	Pt 9·8	Au 2·06	Hg 93·7	Tl 15·9	Pb 19·5	Bi 109	Po	At	Rn
Fr	Ra	Ac															
				Ce 72·6	Pr 63·5	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
				Th	Pa	U	Np	Pu	Am	Cm							

Table VII

Magnetic Susceptibilities ($\chi_A \times 10^6$) at Room Temperature

																H -4	He -2
Li +25											Be -9	B -8	C diamond; -0·5 graphite; -3·5	N -12	O +1700	F	Ne -7
Na +16											Mg +6	Al +16	Si -4	P -27	S -15	Cl -20	A -19
K +22	Ca +44	Sc +315	Ti +150	V +230	Cr +150	Mn +490	Fe ferromagnetics	Co	Ni	Cu -6	Zn -10	Ga -17	Ge -9	As -55	Se -265	Br -32	Kr -28
Rb +19	Sr +92	Y +191	Zr +120	Nb +120	Mo +54	Tc	Ru +43	Rh +110	Pd +550	Ag -22	Cd -20	In -13	Sn α; -42 β; +3	Sb -107	Te -40	I -46	Xe -42
Cs +30	Ba +20	La +140	Hf	Ta +147	W +41	Re +68	Os +9	Ir +25	Pt +156	Au -30	Hg -34	Tl -49	Pb -25	Bi -280	Po	At	Rn
Fr	Ra	Ac															

Ce 2300	Pr ~4000	Nd 5600	Pm	Sm 1820	Eu 30,400	Gd 75,500	Tb 115,000	Dy 102,000	Ho 68,200	Er 44,500	Tm 25,600	Yb 250	Lu

Th +130	Pa	U +620	Np	Pu	Am	Cm

APPENDIX 2

SEMINAR PROBLEMS

1. Why are there only three types of chemical binding?

2. Why do triatomic molecules have different configurations viz: H_2O—a triangle; CO_2—symmetrically arranged on a straight line; N_2O—unsymmetrically arranged on a straight line?

3. Why does the BF_3 molecule have the shape of a plane equilateral triangle while that of NH_3 is a pyramid?

4. Why do some monovalent elements (hydrogen, the halides) form diatomic gases while others (the alkali metals, copper, silver and gold) form metal crystals?

5. Why do silicon and germanium have only one crystalline modification of the diamond type while their homologue carbon has two modifications—diamond and graphite?

6. What is the reason for the great difference in the properties of the two modifications of carbon?

7. What is the reason for the great difference between the chemistries of carbon and silicon? (Carbon, as is well known, is the basis of the organic compounds from the 'animal and vegetable kingdoms', silicon is the basis of the silicates from the 'mineral kingdom'.)

8. Why do the melting points of the halide crystals increase with increase of atomic number while the melting points of the alkali metals fall with increasing atomic number?

9. Why are the elements of the second period, Li, Be, B, C, solids with increasing melting points at room temperature while the following elements N, O, F are gases?

10. Why are things different in the third period? Phosphorus and sulphur, the homologues of nitrogen and oxygen, are solids while chlorine the homologue of fluorine is still a gas.

11. What is the reason for the 'anomalies' in the density of water and ice?

12. Why in the series of homologous compounds H_2O, H_2S, H_2Se, H_2Te does water have 'anomalously' high boiling and freezing points? The differences are respectively of the order of 150° C and 100° C.

13. Why does 'heavy' ice melt at a higher temperature than 'light' ice?

14. Why are the unit cells of metallic hydrides somewhat larger than those of the corresponding deuterides?

15. Why do the melting points of the metals in the fourth period first increase from potassium to chromium, then beginning with iron fall markedly?

16. What is the reason for the large difference in the melting points of the two neighbouring elements: gold (1063° C) and mercury ($-39°$ C)?

17. What substance remains liquid at absolute zero? How can it be crystallized?

18. Why does bismuth contract on melting?

19. Why when the graphite crystal is heated, does it expand along the principal axis and contract in the perpendicular directions while, in contrast, the tellurium crystal under the same conditions contracts along the principal axis and expands in the perpendicular directions?

20. Why is the electrical resistivity of copper several times less than that of the neighbouring elements, nickel and zinc?

21. Why are there such great differences in the electrical properties of homologous elements: diamond is an insulator; silicon and germanium are semiconductors; white tin and lead are metals which can be made superconducting?

22. Why do impurity atoms, even in large (10% or more) amounts, make comparatively little difference to the electrical resistivity of metals while in semiconductors 0·01% of impurity may change the conductivity by a factor of 10^6?

23. Why does a rock salt break under a load of about 1 kg/mm^2 while the calculated value of its strength is 200 kg/mm^2?

24. Is it theoretically possible to produce a material with a strength of 1 tonne/mm^2?

25. What is the physical nature of the high elasticity of rubber which can be extended elastically to 10 times its natural length?

26. Is it possible to make a ferromagnetic alloy from non-ferromagnetic components? Could one make a superconducting compound from non-superconducting components? What are the principles on which such compounds are made?

27. What is the nature of the high values of the dielectric constants of ferroelectrics? (ε may reach values of the order of tens of thousands.)

APPENDIX 3

SOLUTIONS TO THE PROBLEMS

Chapter 2

1. The filled p-subgroup possesses radial symmetry. In fact

$$\psi_{np_x} = f_n(r)\sin\theta\cos\phi;\quad \psi_{np_y} = f_n(r)\sin\theta\sin\phi;\quad \psi_{np_z} = f_n(r)\cos\theta;$$

where $f_n(r)$ is the radial part of the wave function. The total electron density is

$$\rho_{np}^{(r,\theta,\phi)} = \sum_i |\psi_{np_i}|^2 = f_n^2(r)$$

and depends only on r.

2. The table would have a trapezoidal form with a sequence of periods 2, 8, 32, 50. The first four periods would be completely filled. In the fifth period the 5s, 5p, 5d and 5f-subgroups would be filled. The 9 cells of the 5g-subgroup in atoms of element $Z = 100$ would have 8 electrons. The total spin moment of these atoms is $S = 8M_B$, and the total orbital moment $L = 4M_B$. The resulting moment $I = |S-L| = 4M_B$.

Chapter 3

1. $E_1 = \frac{3}{4}I = 10{\cdot}1$ eV, where I is the ionisation potential of the hydrogen atom.

2. (a) $\frac{1}{2}mv^2 = E_1 = 2{\cdot}6$ eV $= 4{\cdot}2\times10^{-12}$ erg; $m = Mm_H = 53{\cdot}2\times10^{-24}$ g.; $v = \sqrt{2E_1/m} = 4$ km/sec; (b) $T = 2E_1/3k = 20{,}000°$ K.

Chapter 4

1. Let A_1 and A_2 be two protons a distance R from one another (Fig. I), B an electron moving in a circular orbit of radius r which is such that the electron is the

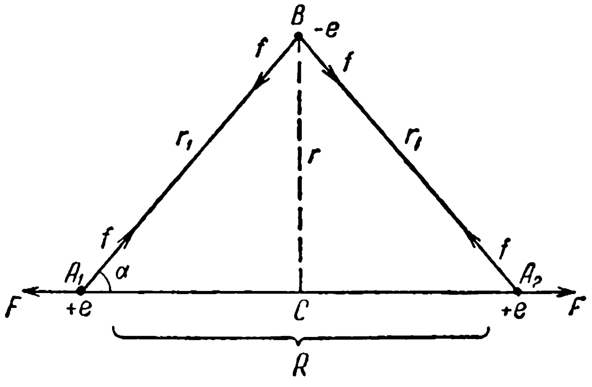


FIG. I

same distance r_1 from both nuclei. The Coulomb attraction force between the electron and each of the nuclei is f, and the Coulomb repulsion force between the nuclei is F:

$$f = -e^2/r_1^2 = -4e^2 \cos^2 \alpha/R^2; \; F = e^2/R^2, \qquad \ldots [\mathit{1}]$$

where r_1 and R are connected by the relationship

$$\tfrac{1}{2}R = r_1 \cos \alpha. \qquad \ldots [\mathit{2}]$$

The equilibrium condition for the nuclei takes the form

$$F = f \cos \alpha. \qquad \ldots [\mathit{3}]$$

From [*1*] and [*2*] we find

$$R^3 = 2r_1^3; \text{ hence } R = r_1\sqrt[3]{2} = 1{\cdot}26r_1. \qquad \ldots [\mathit{4}]$$

From [*2*] and [*4*] we obtain

$$\cos \alpha = R/2r_1 = 2^{-2/3} = 0{\cdot}63; \quad \alpha = 51°. \qquad \ldots [\mathit{5}]$$

The condition of dynamical equilibrium for the electron is given by equating the centripetal force to the Coulomb attraction:

$$mv^2/r = 2f \sin \alpha. \qquad \ldots [\mathit{6}]$$

Since $r = r_1 \sin \alpha$, we have from [*6*]

$$mv^2r = 2e^2 \sin^3 \alpha = 0{\cdot}94e^2 \qquad \ldots [\mathit{7}]$$

(for the hydrogen atom the analogous relationship has the form $mv_H^2 r_H = e^2$). We now introduce the Bohr quantisation condition

$$mvr = nh/2\pi. \qquad \ldots [\mathit{8}]$$

Dividing [*7*] by [*8*], we obtain the orbital velocity of the electron,

$$v = 4\pi e^2 \sin^3 \alpha/nh = v_H \, 2 \sin^2 \alpha = 0{\cdot}94v_H. \qquad \ldots [\mathit{9}]$$

Introducing [*9*] into [*8*], we find for the radius of the molecular orbit

$$r = n^2h^2/8\pi^2me^2 \sin^3 \alpha = r_H/2 \sin^3 \alpha = 1{\cdot}06r_H. \qquad \ldots [\mathit{10}]$$

The equilibrium distance between the nuclei is

$$R = 2r \cot \alpha = (\cot \alpha/\sin^3 \alpha)r_H = 1{\cdot}72r_H = 0{\cdot}91 \text{ Å}. \qquad \ldots [\mathit{11}]$$

The kinetic energy of the electron is

$$K = \tfrac{1}{2}mv^2 = \tfrac{1}{2}mv_H^2 \,.\, 4 \sin^6 \alpha = K_H \,.\, 4 \sin^6 \alpha = 0{\cdot}88K_H. \qquad \ldots [\mathit{12}]$$

The potential energy of the molecular ion is

$$U = -2e^2/r_1 + e^2/R = -(e^2/2r_1)(4 \cos \alpha - 1)/\cos \alpha =$$

$$= \frac{(4e^2 \sin {}^6\alpha)}{r_H} \cdot \frac{(\cos\alpha - \frac{1}{4})}{\sin^2 \alpha \cos \alpha} \qquad \ldots [\mathit{13}]$$

The trigonometrical factor

$$(\cos \alpha - 1/4)/(1 - \cos^2 \alpha) \cos \alpha = (2^{-2/3} - 2^{-2})/(1 - 2^{-4/3}) \, 2^{-2/3} = 1.$$

Consequently,

$$U = -(e^2/r_H)\,4\sin^6\alpha = U_H\,.\,4\sin^6\alpha = 0{\cdot}88U_H; \qquad \ldots [14]$$

(it is worth noting that since the hydrogen atom obeys the virial theorem† $K_H = -\frac{1}{2}U_H$, so also does the molecular ion H_2^+).

The total energy of the H_2^+ ion

$$E = \tfrac{1}{2}U = 0{\cdot}88\;E_H \qquad \ldots [15]$$

is less in absolute magnitude than the energy of the hydrogen atom. So in the Bohr approximation the ion H_2^+ is unstable and must dissociate into a hydrogen atom and a proton.

2. Suppose that both electrons are moving in the same circular orbit and are diametrically opposite to one another. We shall use the same symbols as in the preceding problem for the parameters of the molecule H_2. We write the equilibrium condition for the nuclei thus:

$$e^2/R = (2e^2/r_1^2)\cos\alpha = (e^2/r_1^2)\,.\,(R/r_1), \qquad \ldots [1]$$

whence

$$R = r_1 \text{ and } \alpha = 60°; \qquad \ldots [2]$$

consequently the shape of the molecule corresponds to a regular rhombus turning about an axis passing through the nucleus. The equilibrium condition for each electron, taking into account the force of repulsion between the electrons, is:

$$mv^2/r + e^2/4r^2 = (2e^2/r_1^2)\sin\alpha. \qquad \ldots [3]$$

Hence

$$mv^2r = (e^2/4)(3^{3/2-1}-1) = 1{\cdot}05e^2. \qquad \ldots [4]$$

Using the Bohr quantisation condition (see [*8*] in the preceding problem), we obtain

$$v = (2\pi e^2/nh)\,.\,(3^{3/2}-1)/4 = v_H\,.\,(3^{3/2}-1)/4 = 1{\cdot}05v_H. \qquad \ldots [5]$$

From [*5*], together with [*8*] in the preceding problem, the radius of the molecular orbit is

$$r = nh/2\pi mv = r_H\,4/(3^{3/2}-1) = 0{\cdot}95r_H. \qquad \ldots [6]$$

The internuclear distance is

$$R = (2/\sqrt{3})r = (8/3^{1/2}(3^{3/2}-1))r_H = 1{\cdot}09r_H = 0{\cdot}58\text{ Å}. \qquad \ldots [7]$$

The kinetic energy of the two electrons is

$$K = mv^2 = (e^2/4)(3^{3/2}-1) = (e^2/r_H)\,.\,(3^{3/2}-1)^2/16 = (e^2/r_H)\times 1{\cdot}10. \qquad \ldots [8]$$

---

† As is shown in mechanics, if the potential energy is a power function of the distance between the particles, $U = U(r^s)$, then it is related to the kinetic energy by

$$s\overline{U} = 2\overline{T},$$

where the bar denotes mean value. This relationship is known as the virial theorem. From the virial theorem it follows, in particular, that if the particles interact according to

(*a*) Coulomb's law, then $s = -1$; $\overline{K} = -\frac{1}{2}\overline{U}$; $E = \frac{1}{2}\overline{U}$,

(*b*) Hooke's law, then $s = 2$; $\overline{K} = \overline{U}$; $E = 2\overline{U}$.

The potential energy of the molecule is

$$U = -4e^2/r_1 + e^2/R + e^2/2r = (e^2/R)(3-3^{-\frac{1}{2}}) = -(e^2/8r_H)(3-3^{-\frac{1}{2}})(3^2-3^{+\frac{1}{2}}).$$

Hence we obtain

$$U = -(2e^2/r_H)(3^{3/2}-1)^2/16. \qquad \ldots [9]$$

Equations [*8*] and [*9*] show that $K = -\frac{1}{2}U$, which is a consequence of the virial theorem. The total energy of the molecule

$$E = \tfrac{1}{2}U = -(e^2/r_H)(3^{3/2}-1)^2/16 = -(e^2/r_H)\times 1{\cdot}10 \qquad \ldots [10]$$

is 10% greater in absolute magnitude than the total energy of the two hydrogen atoms

$$2E_H = U_H = -e^2/r_H.$$

So the H_2 molecule is stable in the Bohr approximation. Its binding energy is

$$\Delta E = E - 2E_H = 0{\cdot}2\,I = 2{\cdot}7 \text{ eV}.$$

The results of the calculation for the H_2 molecule are in qualitative agreement with experiment, but quantitatively there is a considerable difference.

$$(\Delta E_{\text{exp}} = 4{\cdot}7 \text{ eV};\ R_{\text{exp}} = 0{\cdot}74 \text{ Å}).$$

Chapter 5

1. (*a*) $Z = 2$; (*b*) $Z = 4$.

2. $q = \pi/2\sqrt{3} = 0{\cdot}91$.

Chapter 6

1. (*a*) $R = a/\sqrt{2}$; (*b*) $R = a\sqrt{3}/4$.

2. $\rho = ZMm_H/v_a$, where Z is the number of particles in the formula, v_a is the volume of the unit cell, M is the molecular weight, and m_H the mass of the hydrogen atom.

Chapter 7

1. (*a*) $\gamma = (1/4)(a/r)^3 = \sqrt{2}/2 = 0{\cdot}71$; (*b*) $\gamma = \frac{1}{2}(a/r)^3 = 4/3\sqrt{3} = 0{\cdot}77$; (*c*) $\gamma = (1/8)(a/r)^3 = 8/3\sqrt{3} = 1{\cdot}54$; (*d*) $\gamma = (1/4)(a/r)^3 = 2$.

2. The series $A = \sum_{i=1}^{\infty} v_i/r_i$ converges rapidly if the summation is taken not over the coordination spheres but over electrically neutral volumes of the crystal (Evjen's method), that is over volumes which contain equal numbers of positive and negative charges. The smallest of them is the unit cell. The next in size will be a cube with edge twice the unit and containing 8 cells. In the structure of CsCl the ions are arranged at the nodes of a body-centred cubic lattice. The coordinates of the nodes [[*hkl*]] of this lattice are all simultaneously either even or odd integers. The even nodes contain ions of one sign, and the odd nodes ions of the other sign. The distances of these nodes from the node taken as origin are given by the relationship $r_i^2 = \sum_{i=1}^{3} h_i^2$. The number v_i of nodes situated at the same distance from the origin is equal to the recurrence factor p, which is defined as the number of nodes which are equivalent from the point of view of symmetry. The coordination numbers v_i' relative to the electrically neutral volume are equal to the v_i for the nodes lying inside the volume. For nodes lying on the faces, edges or at the vertices of the volume, they are respectively 2, 4 and 8 times less. These numbers satisfy the relationship $1+\sum v_i' = 2n^3$, where $n = 1, 2 \ldots$ is the length of the edge of the unit cube in the CsCl lattice.

Below are the values of v_i' for a cube of side $2a$ (ions with coordinates [[311]] lie outside this cube).

$h_ik_il_i$	111	200	220	311	222
$r_i^2 = \Sigma h_i^2$	3	4	8	11	12
$v_i = p_i$	−8	6	12	−24	8
v_i'	−8	3	3	0	1

If in calculating the Madelung constant we limit ourselves to the first volume, for which $n = 1$, then the corresponding series consists of a single term

$$A_{\mathrm{CsCl}} = -(1/8)\times 8/\sqrt{3} = -0{\cdot}58.$$

For the volume with $n = 2$ we obtain

$$A_{\mathrm{CsCl}} = -8/\sqrt{3}+3/\sqrt{4}+3/\sqrt{8}+1/\sqrt{12} = -1{\cdot}77,$$

which is already near enough to the correct value.

3. The energy of formation of ions from atoms is the algebraic sum of the ionisation potential and the electron affinity:

$$\mathrm{Na}+\mathrm{Cl}\begin{cases}\mathrm{Na}^+ + \mathrm{Cl}^- = -5{\cdot}1+\ 3{\cdot}7 = -\ 1{\cdot}4\,\mathrm{eV}.\\ \mathrm{Na}^- + \mathrm{Cl}^+ = +0{\cdot}2-13{\cdot}0 = -12{\cdot}8\,\mathrm{eV}.\end{cases}$$

When the ions approach to a distance of 2·5 Å, which is the interionic distance for the NaCl molecule found spectroscopically, the Coulomb energy of interaction is the same in both cases.

$$U = eV = e\frac{e}{r} = e(4{\cdot}8\times 10^{-10})/(2{\cdot}5\times 10^{-8}) = e\times 1{\cdot}9\times 10^{-2}\ \text{c.g.s. units} = 5{\cdot}7\,\mathrm{eV}.$$

This energy is greater than the work done to form the ions in the first case and is considerably less than it in the second. Thus the ionic molecule is unstable with respect to disintegration into neutral atoms. The energy of dissociation of the ionic molecule $\mathrm{Na^+Cl^-}$ into neutral atoms is 4·6 eV.

Chapter 9

1. In order to calculate the elastic coefficient we shall assume that the electric field is directed perpendicular to the plane of the orbit. As a result of the action of the field the plane of the orbit will be displaced a distance $\bar{x}$ relative to the nucleus and the distance of the electron from the nucleus will be increased to $r = \sqrt{a_0^2+\bar{x}^2}$. The restoring force f is equal to the projection of the Coulomb force on the direction of the field:

$$f = F\cos\alpha = (e^2/r^2)\times(\bar{x}/r) = (e^2/r^2)\bar{x},$$

whence

$$k = e^2/r^3 = e^2/\sqrt[3]{a_0^2+\bar{x}^2}.$$

Since in ordinary fields the deformation $\bar{x}$ of the hydrogen atom is small compared to the radius of the orbit a_0, it follows that to a sufficient degree of accuracy $k \approx e^2/a_0^3$. Introducing this into the value of α given by [*9.6*], we obtain

$$\alpha_H \approx a_0^3 = 0{\cdot}15\times10^{-24}\,\text{cm}^3.$$

2. The equation of motion for a bound electron is $m\ddot{x}+\omega_i^2 x = eE_0 e^{-i\omega t}$. The solution of this equation is

$$x = \frac{eE_0\cos\omega t}{m(\omega_i^2-\omega_0^2)}.$$

The variable dipole moment associated with the forced oscillations of the i-th electron is

$$p_i = ex = \frac{e^2}{m}\cdot\frac{1}{(\omega_i^2-\omega^2)}\cdot E_0\cos\omega t.$$

Hence the contribution of this electron to the polarisability is

$$\alpha_i = \frac{p_i}{E_0\cos\omega t} = \frac{e^2}{m}\cdot\frac{1}{(\omega_i^2-\omega^2)}.$$

3. The surface charge density σ is $P\cos\theta$; the area of the annular zone dS is $2\pi r^2\sin\theta d\theta$. From considerations of symmetry the resulting field must be parallel to the polarisation vector, and so we take the sum of the components of the elementary depolarising fields along $\bar{P}$. In this case

$$E_L = \int_0^\pi \frac{P}{r^2}\cos^2\theta\, 2\pi r^2\sin\theta d\theta = \tfrac{4}{3}\pi P.$$

4. We choose our Z-axis in the direction of the polarisation and find the Z-components of the fields of the elementary dipoles at an arbitrary point of the system chosen as origin. The position of the j-th dipole is given by the radius vector r_j, forming an angle ϕ_j with the Z-axis. The potential of the dipole is

$$v_j = e\left(\frac{1}{r_j}-\frac{1}{r'_j}\right) = \frac{e(r'_j-r_j)}{r_j r'_j} \approx \frac{el_j\cos\phi_i}{r_j^2} = \frac{p_j\cos\phi_j}{r_j^2}.$$

The radial and tangential components of the field are respectively

$$E_j^r = -\frac{\partial v_j}{\partial r_j} = \frac{2p_j \cos\phi}{r_j^3}; \quad E_j^t = -\frac{\partial v_j}{\partial t_j} = -\frac{1}{r_j}\frac{\partial v_j}{\partial \phi_j} = \frac{p_j \sin\phi_j}{r_j^3},$$

the Z-component of the field is

$$E_j^z = E_j^r \cos\phi_j - E_j^t \sin\phi_j = \frac{p_j}{r_j^3}(2\cos^2\phi_j - \sin^2\phi_i) = \frac{p_j}{r_j^3}(3\cos^2\phi_j - 1).$$

The resulting field of the system of dipoles is

$$E_z = \sum_j \frac{p_j}{r_j^3}(3\cos^2\phi_j - 1) = \sum_j \frac{3(\mathbf{p}_j\mathbf{r}_j)^2 - p_j^2 r_j^2}{p_j r_j^5}. \qquad \ldots [1]$$

From considerations of symmetry and uniformity $E_x = E_y = 0$.

5. The points of the cubic complex are obtained by reflections in the planes of symmetry and rotations about the axes of symmetry of the cube. The coordinates $\langle\langle mnp \rangle\rangle$ of the points in the complex are found from the coordinates [[mnp]] of an arbitrarily chosen initial point by carrying out all possible permutations of the indices and of their signs. Permutation and sign-changing of the first two indices give the coordinates of 8 points:

$$\left.\begin{matrix} mnp, & \bar{m}np, & m\bar{n}p, & \bar{m}\bar{n}p \\ nmp, & \bar{n}mp, & n\bar{m}p, & \bar{n}\bar{m}p \end{matrix}\right\}. \qquad \ldots [1]$$

Double cyclic permutation of the indices increases the number of points to 24. Changing all the signs exhausts all the possible combinations and shows that the general cubic complex has multiplicity 48. All the points of the complex have $p_i =$ const and lie on a single coordination sphere $r_i^2 = m^2 + n^2 + p^2 =$ const.

We take an arbitrary straight line with indices [hkl] and a straight line [mnp] passing through the point [[mnp]] of the complex. The cosine of the angle between these straight lines is given by the formula

$$\cos\phi_{mnp} = \frac{hm + kn + lp}{\sqrt{h^2+k^2+l^2}\,\sqrt{m^2+n^2+p^2}}. \qquad \ldots [2]$$

The values of the numerator in [2] corresponding to the points [1] are

$$\left.\begin{matrix} hm+kn+lp, & -hm+kn+lp, & hm-kn+lp, & -hm-kn+lp \\ hn+km+lp, & hn-km+lp, & -hn+km+lp, & -hn-km+lp \end{matrix}\right\} \qquad \ldots [3]$$

The sum of the squares of these numbers is shown in the first line:

$$\left.\begin{matrix} mnp & 4(h^2m^2 + k^2n^2 + h^2n^2 + k^2m^2 + 2l^2p^2) \\ npm & 4(h^2n^2 + k^2p^2 + h^2p^2 + k^2n^2 + 2l^2m^2) \\ pmn & 4(h^2p^2 + k^2m^2 + h^2m^2 + k^2p^2 + 2l^2n^2) \end{matrix}\right\} \qquad \ldots [4]$$

The last two lines are obtained by cyclic permutation of the coordinates of the points. The sum of all the terms in [4] gives

$$8(h^2m^2 + h^2n^2 + h^2p^2 + k^2m^2 + k^2n^2 + k^2p^2 + l^2m^2 + l^2n^2 + l^2p^2). \qquad \ldots [5]$$

The expression in brackets is equal to the square of the denominator in [*2*]. Thus

$$\sum_{i=1}^{24} \cos^2 \phi_{mnp} = 8.$$

Thus for a cubic complex, independently of the direction of orientation of the dipoles,

$$(\cos^2 \phi_{mnp})_{\text{mean}} = \tfrac{1}{3}. \qquad \ldots [6]$$

For points with cubic symmetry in cubic crystals, all the coordination spheres have cubic symmetry and relation [*6*] holds for them. Consequently, it also holds for any set of cubic complexes. For the corresponding points of a cubic crystal in general $(\cos^2 \phi_i)_{\text{mean}} = 1/3$. From equation [*1*] of the solution to problem 4, $E_z = 0$.

6. In electrostatic units the applied potential difference will be

$$V_1 = V = 1000/300 = 3{\cdot}33 \text{ c.g.s. units.}$$

From the piezoelectric tensor for quartz we have

$$r_{11} = d_{11}V = 6{\cdot}8 \times 3{\cdot}33 \times 10^{-8} \text{ cm} = 22{\cdot}6 \text{ Å}; \; r_{22} = -r_{11} = -22{\cdot}6 \text{ Å};$$
$$r_{23} = d_{14} \times V_1 = -1{\cdot}8 \times 3{\cdot}33 \times 10^{-8} = 0{\cdot}013''.$$

7. In point group 2 there is a twofold rotational axis, the direction of which we shall take as the X_3-axis. A rotation of the crystal through 180° about the X_3-axis is described by the following scheme which gives the cosines of the angles between the coordinate axes in their initial and final positions:

	x_1^1	x_2^1	x_3^1
x_1	-1	0	0
x_2	0	-1	0
x_3	0	0	1

The components of the tensor must be the same before and after the rotation: $d'_{ikl} = d_{ikl}$. Consider now whether the component d_{111} vanishes. From considerations of symmetry $d'_{111} = d_{111}$. We apply the formulae for transforming the components of a tensor,

$$d'_{ikl} = c_{mi}c_{nk}c_{ol}d_{mno}$$

$$\begin{aligned}
d'_{111} = {} & c_{11}c_{11}c_{11}d_{111} + c_{11}c_{21}c_{11}d_{121} + c_{11}c_{31}c_{11}d_{131} + \\
& + c_{11}c_{11}c_{21}d_{112} + c_{11}c_{21}c_{21}d_{122} + c_{11}c_{31}c_{21}d_{132} + \\
& + c_{11}c_{11}c_{31}d_{112} + c_{11}c_{21}c_{31}d_{123} + c_{11}c_{31}c_{31}d_{133} + \\
& + c_{21}c_{11}c_{11}d_{211} + c_{21}c_{21}c_{11}d_{221} + c_{21}c_{31}c_{11}d_{231} + \\
& + c_{21}c_{11}c_{21}d_{212} + c_{21}c_{21}c_{21}d_{222} + c_{21}c_{31}c_{21}d_{232} + \\
& + c_{21}c_{11}c_{31}d_{213} + c_{21}c_{21}c_{31}d_{223} + c_{21}c_{31}c_{31}d_{233} + \\
& + c_{31}c_{11}c_{11}d_{311} + c_{31}c_{21}c_{11}d_{321} + c_{31}c_{31}c_{11}d_{331} + \\
& + c_{31}c_{11}c_{21}d_{312} + c_{31}c_{21}c_{21}d_{322} + c_{31}c_{31}c_{21}d_{332} + \\
& + c_{31}c_{11}c_{31}d_{313} + c_{31}c_{21}c_{31}d_{323} + c_{31}c_{31}c_{31}d_{333}.
\end{aligned}$$

From the scheme of cosines it follows that

$$c_{11} = c_{22} = -c_{33} = -1; \quad c_{12} = c_{21} = c_{23} = c_{32} = c_{31} = c_{13} = 0.$$

Substituting the values of the cosines in the transformation formula, we obtain $d'_{111} = -d_{111}$. This relationship taken with the one above shows that $d'_{111} = d_{111} = 0$. From the transformation formulae we obtain by a similar method the result that for crystals with symmetry 2 a number of the other components of the tensor vanish and the piezoelectric tensor takes the form

	t_{11}	t_{22}	t_{33}	t_{23}	t_{31}	t_{12}
p_x	0	0	0	d_{14}	d_{15}	0
p_y	0	0	0	d_{24}	d_{25}	0
p_z	d_{31}	d_{32}	d_{33}	0	0	d_{36}

Chapter 11

1. From [*11.6*] with $c = 1$

$$u_{AA} = \frac{2L}{zN} = \frac{2\times 81{\cdot}2}{12}\ \text{kcal/mole} = 13{\cdot}5\ \text{kcal/mole} = \frac{13{\cdot}5}{23{\cdot}4}\ \text{eV} \cong 0{\cdot}6\ \text{eV}.$$

2. The maximum solubility temperature T_m is found from [*11.11*] with $c = 0{\cdot}5$· Taking logarithms in [*11.11*] gives $\ln(1-c)/c = p(1-2c)$, where $p = zU_m/kT$. From the equation

$$p = \left(\frac{\ln\frac{1-c}{c}}{1-2c}\right)_{c=1/2} = \frac{0}{0} = \frac{-\frac{c}{1-c}\cdot\frac{1}{c^2}}{-2} = \left[\frac{1}{2c(1-c)}\right]_{c=1/2} = 2.$$

Consequently,

$$T_m = \frac{2zU}{k};\ U_m = \frac{kT_m}{2z} = \frac{RT_m}{24} = \frac{1100}{12}\ \text{kcal/mole} \approx 0{\cdot}1\ \text{kcal/mole} \approx 0{\cdot}004\ \text{eV}.$$

The energy of mixing, which is equal to the difference in energy between the heterogeneous and homogeneous bonds, is thus less than 1% of the energy of either type of bond.

Chapter 12

1. (*a*) We cut the pseudocubic cell AuCu into halves each of which is body-centred tetragonal with $a_t\sqrt{2} = a_c$, $c_t = a_c$, where a_c is the length of the edge of the initial cell. We denote the sites at the centres of the cells by α, and those at the vertices by β. When they are arranged in order the coordination number $z = 12$ splits into two subgroups $8+4$

$$n_{AA} = N_\alpha p_\alpha^A z_\alpha^\alpha p_\alpha^A \tfrac{1}{2} + N_\beta p_\beta^A z_\beta^\beta p_\beta^A \tfrac{1}{2} + N_\alpha p_\alpha^A z_\alpha^\beta p_\beta^A =$$

$$= \tfrac{1}{2}\,Np4p\,\tfrac{1}{2} + \tfrac{1}{2}\,No4o\,\tfrac{1}{2} + \tfrac{1}{2}\,Np8o = N(1^2+o^2+4po) = N(1+2po).$$

$$n_{AB} = N_\alpha p_\alpha^A z_\alpha^\alpha p_\alpha^B + N_\beta p_\beta^B p_\beta^A + N_\alpha p_\alpha^A z_\alpha^\beta p_\beta^B + N_\alpha p_\alpha^B z_\alpha^\beta p_\beta^A =$$

$$= \tfrac{1}{2}\,\mathrm{N}p4o + \tfrac{1}{2}\,Np4o + \tfrac{1}{2}\,Np8p + \tfrac{1}{2}\,No8o = 4N(p^2+o^2-po) = 4N(1-po),$$

where

$$p = p_\alpha^A = p_\beta^B; \; o = p_\alpha^B = p_\beta^A = 1-p.$$

Introducing in place of p the degree of order s, we obtain

$$n_{AA} = n_{BB} = \tfrac{1}{12} n(L-s^2); \quad n_{AB} = \tfrac{1}{6} n(3+s^2).$$

$$(b) \quad n_{AA} = N_\alpha^A (8p_\alpha^A \tfrac{1}{2} + 4p_\beta^A) = 3Np_\alpha^A (p_\alpha^A + p_\beta^A);$$

$$n_{BB} = N_\alpha^B (8p_\alpha^B \tfrac{1}{2} + 4p_\beta^B) = 3Np_\alpha^B (p_\alpha^B + p_\beta^B);$$

$$n_{AB} = N_\alpha^A (8p_\alpha^B + 4p_\beta^B) + N_\alpha^B 4p_\beta^A = 3N[p_\alpha^A (2p_\alpha^B + p_\beta^B) + p_\alpha^B p_\beta^A].$$

The probabilities are expressed in terms of the degree of order as follows, for $c_A = 3/4$:

$$p_\alpha^A = \tfrac{1}{4}(3+s); \quad p_\alpha^B = \tfrac{1}{4}(1-s); \quad p_\beta^B = \tfrac{1}{4}(1+3s); \quad p_\beta^A = \tfrac{3}{4}(1-s).$$

Consequently,

$$n_{AA} = \tfrac{1}{16} n(9-s^2); \quad n_{BB} \tfrac{1}{16} n(1-s^2); \quad n_{AB} = \tfrac{1}{8} n(3+s^2).$$

2. (*a*) $c_A \leqslant \tfrac{1}{2}$. All the β sites are occupied by B atoms while of the α-sites $c_A N$ are occupied by A atoms and $(\tfrac{1}{2}-c_A) N$ by B atoms.

$$n_{AB} = c_A N \times 8 + (\tfrac{1}{2}-c_A) N \times 4 \times 2c_A = \tfrac{2}{3} nc_A (3-2c_A).$$

(*b*) (1) $c_A < 1/4$. A atoms occupy $c_A N$ of the α-sites: $n_{AB} = c_A N \cdot 12 = 12nc_A$;

(2) $c_A > 1/4$. Part of the β-sites, $\Delta N_A = c_A N - (1/4) N = (c_A - 1/4) N$ is occupied by A atoms. The concentration of A atoms in the β-sites is $\Delta N_A/(3/4) N = (1/3)(4c_A - 1)$; the concentration of B atoms, $\dfrac{1-c_A}{\frac{3}{4}N} = \tfrac{3}{4}(1-c_A)$.

$$n_{AB} = \tfrac{1}{4} N \times 12 \times \tfrac{4}{3}(1-c_A) + \tfrac{3}{4} N \times \tfrac{1}{3}(4c_A - 1) \times 12 \times \tfrac{4}{3}(1-c_A).$$

Hence

$$n_{AB} = \tfrac{8}{3} nc_A (1-c_A).$$

3. Consider the concentration range $0 \leqslant c \leqslant \tfrac{1}{2}$.

$$N_\alpha^A = \tfrac{1}{2} N(1+s)c; \quad N_\alpha^B = \tfrac{1}{2} N[1-(1+s)c]; \quad N_\beta^A = \tfrac{1}{2} N(1-s)c;$$

$$N_\beta^B = \tfrac{1}{2} N[1-(1-s)c];$$

$$n_{AA} = 4Nc^2(1-s^2); \quad n_{BB} = 4N[(1-c)^2]; \quad n_{AB} = 8N[c(1-c)+c^2 s^2].$$

The internal energy of the alloy is

$$U = 4N[cU_{AA} + (1-c) U_{BB} + 2c(1-c) U_1] + 8Nc^2 s^2 U_1 + C.$$

The entropy is

$$S = -\tfrac{1}{2} Nk \{(1+s) c \ln (1+s) c + (1-s) c \ln (1-s) c + \\ + [1-(1+s)c] \ln [1-(1+s)] + [1-(1-s)c] \ln [1-(1-s)c]\} + C'.$$

The equilibrium equation is

$$\ln \frac{(1+s)[1-(1-s)c]}{(1-s)[1-(1+s)c]} = \frac{32csU_1}{kT} = \frac{8cT_k}{T} s, \quad \text{where} \quad T_k = -\frac{4U_1}{k}.$$

The gradient of the tangent to the logarithmic curve at the origin is tan $\phi = 2/(1-c)$; and equating it to the gradient of the straight line, we obtain

$$T_k(c) = 4T_k\,(\tfrac{1}{2})\,c\,(1-c).$$

This is the equation of a parabola.

Chapter 13

1. Taking account of viscosity, the rate of formation of nuclei of crystallization in a condensed phase is given by

$$c \sim e^{-\frac{1}{kT}\left\{Q+\frac{1}{3}\pi\sigma\left[\frac{4\sigma T_0}{q(T-T_0)}\right]^2\right\}} = e^{-\frac{f(x)}{kT_0}}, \qquad \ldots [1]$$

where Q is the activation energy for diffusion, σ is the surface tension at the boundary between the nucleus and the mother-phase, T_0 is the temperature of crystallization and q is the latent heat of crystallization. Introducing the new variable $x = T_0/T$, the function in the exponent in formula [*1*] becomes

$$f(x) = x\left[Q+a\left(\frac{x}{x-1}\right)^2\right] = Qx+\frac{ax^3}{(x-1)^2}; \qquad \ldots [2]$$

where

$$a = \frac{\pi\sigma}{3}\left(\frac{4\sigma}{q}\right)^2. \qquad \ldots [3]$$

From the condition for a turning point we have

$$f'(x) = Q+\frac{3ax^2}{(x-1)^2}-\frac{2ax^3}{(x-1)^3} = 0.$$

whence we obtain the equation

$$ax^2\,(3-x) = Q\,(x-1).^3 \qquad \ldots [4]$$

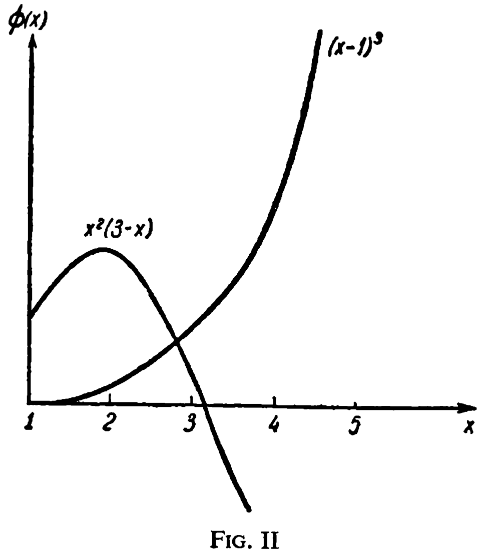


FIG. II

Fig. II shows the graphical solution of equation [*4*] for the case where $Q \simeq a$. The order of magnitude of a is obtained from [*3*]:

$$a \approx \frac{16\sigma^3}{q^3} \approx \tfrac{1}{2}\,\text{eV} \quad (\sigma = 80\ \text{erg/cm}^2;\quad q = 80\ \text{cal/cm}^3).$$

Putting $Q \simeq 2a$ calculation from equation [*4*] gives the result that the minimum of the function $f(x)$ lies near to $x = 2{\cdot}2$. Hence the temperature corresponding to the maximum rate of formation of nuclei is

$$T_m = 0{\cdot}45\,T_0.$$

We may note that the curve $c = c(T)$ has a maximum even if there is no viscosity. Putting $Q = 0$ in [*4*] we obtain $T_m = 0{\cdot}33\,T_0$.

2. (*a*) $q = 80$ cal/cm^3;
$\sigma^{273^\circ} = 75$ erg/cm^2;
$\sigma^{263^\circ} = 76{\cdot}5$ erg/cm^2,
when $\Delta T = 1^\circ$, $r_k = 12\times10^{-6}$ cm
when $\Delta T = 10^\circ$, $r_k = 1{\cdot}2\times10^{-6}$ cm
$\Delta F_k = 450\times10^{-12}$ erg

(*b*) $T_0 = 1083^\circ\,\text{C} = 1353^\circ\,\text{K}$;
$\sigma = 1100$ erg/cm^2;
$q = 51$ cal/gm $= 430$ cal/cm^3,
for $\Delta T = 10^\circ$, $r_k = 17\times10^{-6}$ cm.

3. If collisions between the growing crystals are neglected, the crystallization kinetics are described by the integral equation

$$u(t) = 1-\tfrac{2}{3}p^4\int_0^t u(\tau)(t-\tau)^3\,d\tau, \qquad \ldots [1]$$

where

$$p^4 = 12cw^3. \qquad \ldots [2]$$

Equation [*1*] is a second-order integral equation of the Volterra type

$$\phi(t) = f(t)+\lambda\int_0^t K(t,\tau)\,\phi(\tau)\,d\tau, \qquad \ldots [3]$$

the solution of which is expressed by the series

$$\phi(t) = \phi_0(t)+\lambda\phi_1(t)+\lambda^2\phi_2(t)+\ldots+\lambda^n\phi_n(t)+\ldots, \qquad \ldots [4]$$

where

$$\phi_0(t) = f(t);\quad \phi_{n+1}(t) = \int_0^t K(t,\tau)\,\phi_n(\tau)\,d\tau. \qquad \ldots [5]$$

Applying formula [*5*] to equation [*1*] we obtain

$$\begin{aligned} &u_0(t) = 1 && u_0(t) = 1 \\ &u_1(t) = \frac{6t^4}{4!} && \lambda u_1(t) = -\frac{4p^4t^4}{4!} \\ &u_2(t) = \frac{6^2(t^4)^2}{8!} && \lambda^2 u_2(t) = \frac{(4p^4t^4)^2}{8!} \\ &\ldots\ldots\ldots\ldots\ldots\ldots && \ldots\ldots\ldots\ldots\ldots\ldots \\ &u_n(t) = \frac{6^n(t^4)^n}{(4n)!} && \lambda^n u_n(t) = \frac{(-4p^4t^4)^n}{(4n)!}. \end{aligned} \qquad \ldots [6]$$

and substituting these values from [*6*] into [*4*] we obtain the solution of equation [*1*]

$$u(t) = 1 - \frac{4p^4t^4}{4!} + \frac{(4p^4t^4)^2}{8!} - \ldots + \frac{(-4p^4t^4)^n}{(4n)!} + \ldots \qquad \ldots [7]$$

This solution can be written in a more compact form. Using the series expansions for the hyperbolic and trigonometrical cosines we find that their product is

$$\cosh x \cos x = \left(1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \frac{x^8}{8!} + \ldots\right)\left(1 - \frac{x^2}{2!} - \frac{x^4}{4!} - \frac{x^6}{6!} - \frac{x^8}{8!} - \ldots\right) =$$

$$= 1 - \frac{4x^4}{4!} + \frac{(4x^4)^2}{8!} - \ldots \qquad \ldots [8]$$

Comparing [*7*] and [*8*] shows that equation [*7*] can be written in the form

$$u(t) = \cosh pt \cos pt.$$

4. (*a*) For the two cases we are interested in

$$u_{\text{I}} = \cosh t \cos t; \quad u_{\text{II}} = e^{-t^4/6} \quad (p = 1). \qquad \ldots [1]$$

For ease of calculation it is convenient to measure the time by three different scales: in degrees of arc, $t°$; in radians t_{rad} and in $t' = (2/\pi)\, t_{rad}$. The detailed calculations are given in the table.

$t°$	t_{rad}	$t' = 0{\cdot}638 t_{rad}$	$\cos t°$	$\cosh t'$	u_{I}	t^4_{rad}	$x = \frac{1}{6}t^4$	u_{II}
0	0·00	0·00	1·00	1·00	1·00	0·00	0·00	1·00
30	0·52	0·33	0·87	1·14	0·99	0·08	0·01	0·99
50	0·87	0·56	0·64	1·41	0·90	0·58	0·10	0·90
60	1·04	0·67	0·50	1·60	0·80	1·20	0·20	0·82
70	1·22	0·78	0·34	1·85	0·63	2·22	0·37	0·69
80	1·40	0·89	0·17	2·14	0·37	3·80	0·63	0·53
90	1·57	1·00	0·00	2·51	0·00	6·05	1·01	0·36
100	1·74	—	—	—	—	9·18	1·53	0·22
110	1·92	—	—	—	—	13·6	2·27	0·10
130	2·26	—	—	—	—	26·5	4·4	0·01
330	2·57	—	—	—	—	42	7·0	0·001

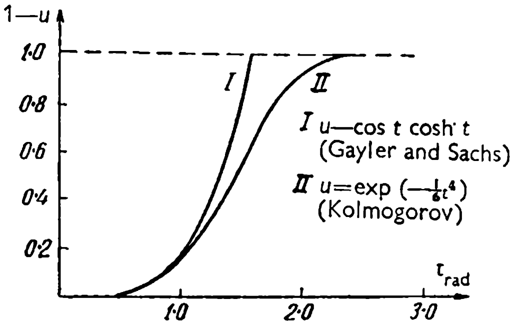


FIG. III

The results of the calculation are given graphically in Fig. III.

(*b*) In case I the time for complete crystallization is $t_r = \pi/2 = 1{\cdot}57$. In case II the process of final crystallization goes on indefinitely. If the end of crystallization is taken to be the time when 0·999 of the total volume has crystallized then $u_{\text{II}} = 0{\cdot}001$ gives $t_r' = 2{\cdot}54$, a time which is 1·6 times greater than in Case I. Beginning at $t = 1{\cdot}04$ (or $0{\cdot}4t_r'$) curve I rises rapidly above curve II. At this time the volume which has crystallized is about 20% of the total.

(*c*) In case I

$$N(t_r) = c\int_0^{t_r} \cos p\tau \cosh p\tau d\tau = \frac{c}{p}\int_0^{\pi/2} \cos x \cosh x dx, \text{ where } x = p\tau. \qquad \dots [2]$$

In the formula for integrating by parts

$$\int uv' = uv - \int u'v,$$

put $u = \cos x$, $v' = \cosh x$ so that $u' = -\sin x$ and $v = \sinh x$ and we obtain

$$\int \cos x \cosh x dx = \cos x \sinh x - \int \sin x \sinh x dx. \qquad \dots [3]$$

If we take $u = \cosh x$; $v' = \cos x$ so that $u' = \sinh x$; $v = \sin x$ we obtain

$$\int \cos x \cosh x dx = \sin x \cosh x + \int \sin x \sinh x dx. \qquad \dots [4]$$

Combining [*3*] and [*4*] gives

$$\int_0^{\pi/2} \cos x \cosh x dx = \tfrac{1}{2}[\cos x \sinh x + \sin x \cosh x]_0^{\pi/2} = 0{\cdot}100.$$

and substituting this value in [*2*] we obtain

$$N(t_r) = 0{\cdot}100\,\frac{c}{p}. \qquad \dots [5]$$

In case II

$$N(t_r) = N(\infty) = c\int_0^{\infty} e^{-p^4\tau^4/6} d\tau. \qquad \dots [6]$$

The value of this integral can be found in Jahnke and Emde's 'Tables of Functions',

$$\int_0^\infty e^{-wt^n}\, t^z dt = \frac{1}{z+1}\,\frac{\left(\frac{z+1}{n}\right)!}{w^{\frac{z+1}{n}}}$$

which with $z = 0$, $n = 4$ and $w = p^4/6$ gives

$$N(\infty) = \frac{\sqrt{6}}{4!}\,\frac{c}{p} = 0{\cdot}065\,\frac{c}{p}. \qquad \ldots [8]$$

5. As in the three-dimensional case the non-crystallizing part of the unit area is

$$u(t) = \exp\,(-c\Omega)$$

where

$$\Omega = \int_0^t s(\tau)\,d\tau = 4w^2 \int_0^t (t-\tau)^2\, d\tau = \tfrac{4}{3}w^2t^3.$$

Putting this in the formula for $u(t)$ we obtain

$$u(t) = \exp\,(-\tfrac{4}{3}\,cw^2t^3)$$

where c is the number of nuclei forming in unit time on unit area of a two-dimensional melt.

6. The kinetics of crystallization are determined by the rate of formation of nuclei c, the linear rate of growth w and the time t. The dimensions of these quantities are: (*a*) Three-dimensional case

$$[c] = L^{-3}T^{-1}; \quad [w] = LT^{-1}; \quad [t] = T.$$

A dimensionless parameter which determines the kinetics of crystallization can be constructed with the form

$$a = c^pw^qt^r = L^{q-3p}T^{r-p-q}.$$

and from it we obtain the two equations

$$q = 3p; \quad r = p+q = 4p.$$

Putting the values of q and r into the equation for α gives

$$\alpha = c^pw^{3p}t^{4p} = (cw^3t^4)^p.$$

This dimensional analysis can be seen to agree with the exact result (see [*13.23*]) which was that the kinetics of crystallization are determined by the product cw^3t^4.

(*b*) Two-dimensional case:

$$[c] = L^{-2}T^{-1}; \quad [w] = LT^{-1}; \quad [t] = T.$$

$$\alpha = c^p w^q t^r = L^{q-2p} T^{r-p-q}.$$

$$q = 2p; \quad r = p+q = 3p$$

$$\alpha = c^p w^{2p} t^{3p} = (cw^2t^3)^p,$$

which agrees with the full calculation given in problem 5 above.

7.

$$\text{(a)} \quad w_r = w_0\left(1-\frac{r_k}{r}\right); \qquad \text{(b)} \quad v_r = 4\pi w_0 \int_0^t \left(1-\frac{r_k}{r}\right) r^2 dr$$

(for a spherical nucleus), where w_0 is the rate of growth of a plane crystalline front.

Chapter 14

1. (***a***) The coefficients we want are characteristic of three basic thermodynamic processes:
volume expansion for an isobaric process

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p, \qquad \ldots [1]$$

volume compressibility for an isothermal process

$$\beta = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T, \qquad \ldots [2]$$

and thermal elasticity for an isochoric process

$$\gamma = \frac{1}{p}\left(\frac{\partial p}{\partial T}\right)_V. \qquad \ldots [3]$$

The parameters p, V and T are connected by the equation of state for the body:

$$f(p, V, T) = 0. \qquad \ldots [4]$$

Hence

$$V = V(p, T). \qquad \ldots [5]$$

Now use the formula for differentiating a composite function

$$Z = Z[x(y), y, u], \qquad \ldots [6]$$

which depends on variables x and y and a parameter u:

$$\left(\frac{\partial Z}{\partial y}\right)_u = \left(\frac{\partial Z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_u + \left(\frac{\partial Z}{\partial y}\right)_x. \qquad \ldots [7]$$

In formula [*6*] put:

$$Z = V; \quad u = V; \quad x = p; \quad y = T,$$

and we obtain from [*7*]

$$\left(\frac{\partial V}{\partial T}\right)_V = 0 = \left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_V + \left(\frac{\partial V}{\partial T}\right)_p. \qquad \ldots [8]$$

Hence

$$\left(\frac{\partial V}{\partial T}\right)_p = -\left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_V. \qquad \ldots [9]$$

Substituting [*1*] to [*3*] into [*9*], gives the required relationship:

$$\alpha = \beta\gamma p. \qquad \ldots [10]$$

(*b*) The equation of state for an ideal gas is

$$pV = RT, \qquad \ldots [11]$$

from which we have

$$\alpha = \frac{1}{T}; \quad \beta = \frac{1}{p}; \quad \gamma = \frac{1}{T}. \qquad \ldots [12]$$

It is easy to see that relationships [*12*] satisfy [*10*].

2. The entropy of the system is

$$dS = \frac{\delta Q}{T} = \frac{1}{T}\,dU + \frac{p}{T}\,dV, \qquad \ldots [1]$$

since from the first law of thermodynamics

$$\delta Q = dU + p dV. \qquad \ldots [2]$$

Using the equation of state ([*4*] in the preceding problem), we can express the internal energy as a function of the two parameters, $U = U(V, T)$. We find the total differential of this function

$$dU = \left(\frac{\partial U}{\partial V}\right)_T dV + \left(\frac{\partial U}{\partial T}\right)_V dT, \qquad \ldots [3]$$

and substitute it in [*1*]:

$$dS = \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V dT + \left[\frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_T + \frac{p}{T}\right] dV. \qquad \ldots [4]$$

The condition for dS to be a total differential is

$$\frac{\partial}{\partial V}\left[\frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V\right]_T = \frac{\partial}{\partial T}\left[\frac{1}{T}\left(\frac{\partial U}{\partial V}\right)_T + \frac{p}{T}\right]_V. \qquad \ldots [5]$$

Carrying out the differentiation, we find

$$\left(\frac{\partial U}{\partial V}\right)_T + p = T\left(\frac{\partial p}{\partial T}\right)_V. \qquad \ldots [6]$$

Putting in formula [7] of the preceding problem

$$Z = U; \quad u = p; \quad x = V; \quad y = T,$$

we find

$$\left(\frac{\partial U}{\partial T}\right)_p = \left(\frac{\partial U}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p + \left(\frac{\partial U}{\partial T}\right)_V. \qquad \ldots [7]$$

Differentiation of the internal energy with respect to the temperature gives

$$\left(\frac{\partial U}{\partial T}\right)_p = \left(\frac{\partial Q}{\partial T}\right)_p - p\left(\frac{\partial V}{\partial T}\right)_p = c_p - \alpha p V \qquad \ldots [8]$$

$$\left(\frac{\partial U}{\partial T}\right)_V = \left(\frac{\partial Q}{\partial T}\right)_V - 0 = c_v. \qquad \ldots [9]$$

Substituting [8] and [9] in [7] and then using [6], we obtain

$$c_p - c_v = \alpha V\left[p + \left(\frac{\partial U}{\partial V}\right)_T\right] = \alpha V T\left(\frac{\partial p}{\partial T}\right)_V = \alpha \gamma p V T.$$

Using [10] of the preceding problem, we obtain finally

$$c_p - c_v = \frac{\alpha^2}{\beta} VT.$$

3. The specific volume of iron $v = A/\rho = 56/7{\cdot}9 = 7{\cdot}1$ cm^3/mole;
$c_p - c_v = 0{\cdot}7$ cal/mole.°K.

4. $$c_v = \frac{R\beta^2 e^\beta}{(e^\beta - 1)^2}, \quad \text{where} \quad \beta = \frac{h\nu}{kT}.$$

5. (*a*) The null expansion condition for a uniaxial crystal has the form

$$\alpha_{11}(c_{13}^2 + c_{23}^2) + \alpha_{33}c_{33}^2 = 0. \qquad \ldots [1]$$

Since $c_{13}^2 + c_{23}^2 + c_{33}^2 = 1$, we have

$$c_{33}^2 = \frac{\alpha_{11}}{\alpha_{11} - \alpha_{33}}. \qquad \ldots [2]$$

(*b*) We denote by θ the angle between the desired direction and the principal axis of symmetry of the crystal. Then $c_{33} = \cos\theta$, and formula [2] can be written in the form

$$\tan^2\theta = -\alpha_{33}/\alpha_{11}. \qquad \ldots [3]$$

Since $\tan\theta$ must be real, it follows that the condition for the existence of a direction of zero expansion is that the coefficients of thermal expansion should have different signs.

(*c*) In the calcite crystal the null expansion directions form a circular cone with internal angle $2\theta = 64° \, 43'$, and with axis parallel to the threefold symmetry axis of the crystal.

6. The volume expansion coefficient β is connected with the linear coefficients by the relationship

$$\alpha_{\parallel} + 2\alpha_{\perp} = \beta. \qquad \ldots [1]$$

Assuming for simplicity that a linear law holds, we obtain

$$\left(\frac{c}{a}\right)_t = \left(\frac{c}{a}\right)_0 \frac{(1+\alpha_{11}\Delta t)}{(1+\alpha_{\perp}\Delta t)} = \left(\frac{c}{a}\right)_0 [1-(\alpha_{\parallel}-\alpha_{\perp})\Delta t].$$

Hence

$$\alpha_{\parallel} - \alpha_{\perp} = \frac{\left(\frac{c}{a}\right)_t - \left(\frac{c}{a}\right)_0}{\left(\frac{c}{a}\right)_0 \Delta t}. \qquad \ldots [2]$$

From equations [*1*] and [*2*] we can determine the required coefficients.

7. The direction of one of the principal axes of the ellipsoid OX_2 is parallel to the monoclinic axis [010] of the crystal. To determine α_{22} we have to measure the thermal expansion along the monoclinic axis. The two other principal axes of the ellipsoid, OX_1 and OX_3, lie in the monoclinic plane (010) (Fig. IV), where *a* and *c* are the axes of the crystal. The angle ψ determines the position of the thermal expansion tensor relative to the crystallographic axes.

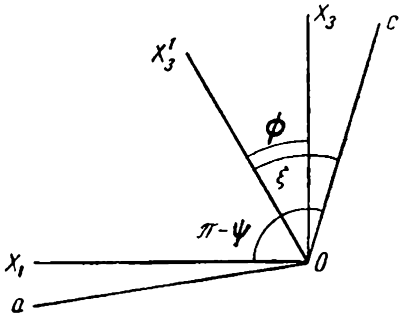


FIG. IV

Take an arbitrary direction OX_3' lying in the monoclinic plane and forming an angle ξ with the *c* axis. The coefficient of linear expansion along this axis, α, is connected with the axes α_{11} and α_{33} of the thermal expansion tensor by the relationship

$$\alpha_{33}' = c_{13}^2\alpha_{11} + c_{33}^2\alpha_{33}. \qquad \ldots [1]$$

Denoting the angle between OX_3' and OX_3 by ϕ, we obtain

$$\alpha_{33}' = \alpha_{11}\sin^2\phi + \alpha_{33}\cos^2\phi = \frac{\alpha_{11}+\alpha_{33}}{2} - \frac{\alpha_{11}-\alpha_{33}}{2}\cos 2\phi. \qquad \ldots [2]$$

Since

$$\phi = (\xi+\psi) - \frac{\pi}{2}, \qquad \ldots [3]$$

it follows that

$$\alpha'_{33} = \frac{\alpha_{11}+\alpha_{33}}{2}+\frac{\alpha_{11}-\alpha_{33}}{2}(\cos 2\xi \cos 2\psi - \sin 2\xi \sin 2\psi). \qquad \ldots [4]$$

We introduce the notation:

$$\tfrac{1}{2}(\alpha_{11}+\alpha_{33}) = A; \quad \tfrac{1}{2}(\alpha_{11}-\alpha_{33})\cos 2\psi = B; \quad \tfrac{1}{2}(\alpha_{11}-\alpha_{33})\sin 2\psi = C. \qquad \ldots [5]$$

Then equation [4] can be written in the form

$$A+B\cos 2\xi + C\sin 2\xi = \alpha'_{33}. \qquad \ldots [6]$$

By measuring the coefficients of linear expansion α'_{33} and the angles ξ for three directions lying in the monoclinic plane, we can obtain three equations corresponding to [6] and from them determine the quantities A, B and C. After that, on solving the system of equations [5], we find the quantities which define the thermal expansion tensor

$$\alpha_{11} = A+\frac{B}{\cos 2\psi}; \quad \alpha_{33} = A-\frac{B}{\cos 2\psi};$$

$$\tan 2\psi = -\frac{C}{B}.$$

Chapter 15

1. Taking logarithms in [15.2], we obtain

$$\frac{Q}{R}\cdot\frac{1}{T} = \ln A - \ln D. \qquad \ldots [1]$$

Equation [1] is a straight line in the variables $1/T$ and $\ln D$. From the gradient of this line we obtain Q, and from the magnitude of the ordinate at the origin we obtain A.

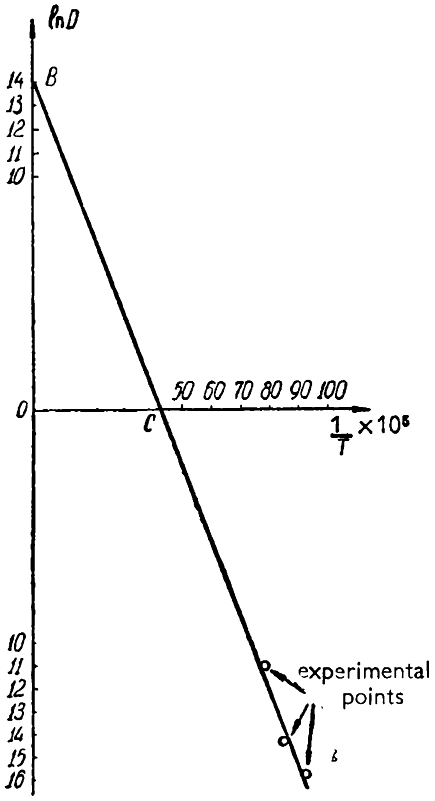


FIG. V

T °K	$1/T$	D cm²/day	log D	log D	ln D
1073	93×10^{-5}	$1\cdot3\times10^{-7}$	$\bar{7}\cdot11$	$-6\cdot89$	$-15\cdot9$
1173	85×10^{-5}	$5\cdot0\times10^{-7}$	$\bar{7}\cdot70$	$-6\cdot30$	$-14\cdot5$
1273	$78\cdot5\times10^{-5}$	$1\cdot4\times10^{-5}$	$\bar{5}\cdot15$	$-4\cdot85$	$-11\cdot2$

In Fig. V the line has been drawn from experimental points (with known error). From the graph

$$\frac{Q}{R} = \frac{OB}{OC} = \frac{13\cdot7\times10^5}{42} = 32\cdot3\times10^3;$$

$$Q = 64\cdot6 \text{ kcal/mole};$$

$$\ln A = (\ln D)_{\frac{1}{T}=0} = OB = 13\cdot7;$$

$$\log A = 0\cdot434\times13\cdot7 = 5\cdot95;$$

$$A = 8\cdot9\times10^5 \text{ cm}^2\text{/day}.$$

2. (*a*)

T °K	$(Q/RT)\times 0{\cdot}43$	D cm²/day.
300	20·5	$1{\cdot}8\times 10^{-15}$
438	14	5×10^{-9}
600	~10	$5{\cdot}8\times 10^{-5}$

(*b*) From the transport equation for a substance in a stationary process we find that the time t necessary for the transport of a mass m is $t = mdx/Dsdc$. Putting $m = 1$, $S = 1$, $dc = 1$, we obtain $t = 5\times 10^7$ days, or $1{\cdot}4\times 10^5$ years. For the non-stationary case the value of the concentration c at a distance x from the boundary of separation at time t is given by the expression†

$$c = \tfrac{1}{2}c_0\left[1-\frac{2}{\sqrt{\pi}}\right]\int\limits_0^{\frac{x}{2\sqrt{Dt}}} e^{-y^2}\,dy.$$

Using the value of the error integral‡, we find from this that after the given interval of time the concentration of radioactive lead at a depth of 0·1 mm is given by $c = 0{\cdot}45$.

---

† R. Barrer, 'Diffusion in and through Solids'. Cambridge U.P. 1951.

‡ H. B. Dwight, *Tables of Integrals and other Mathematical Data*, 4th ed. Macmillan, New York, 1961.

BIBLIOGRAPHY

The bibliography in the Russian edition of this book gave extensive lists of supplementary text books in English, French, German and Russian. Nearly all of them were available in English either as their original language or as translations. These so available have been retained in the list below and a few others added.

References to advanced work, in the Russian edition took the form of extensive references to original papers again in many languages. It has seemed preferable to replace these by references to recent review articles.

References frequently cited

The following short list includes the major review articles in Crystal Physics.

Solid State Physics: Advances in Research and Applications, edited by F. Seitz and D. Turnbull, Academic Press, N.Y. This is cited as *Solid State Physics* with the volume number and date.

Reports on Progress in Physics. Annual volume published by the Institute of Physics and the Physical Society, London. This is cited as *Rep. Progr. Physics* with the volume number and date.

Progress in Metal Physics edited by B. Chalmers and R. King. The title of this series was changed to *Progress in Materials Science* in 1961. It is published by Pergamon Press.

GENERAL READING

BARRETT, C. S. *Structure of Metals.* McGraw Hill, New York and London 1943.
DEKKER, A. J. *Solid State Physics.* Macmillan, London 1958.
JOFFE, A. F. *The Physics of Crystals.* McGraw Hill, New York and London 1928.
KITTELL, C. *Solid State Physics.* Wiley, New York, Chapman and Hall, London 1953.
LANDAU, L. D. and E. M. LIFSHITZ. *Course of Theoretical Physics.* Pergamon Press.
  VOL. I. *Mechanics.* This includes the virial theorem, tensor algebra, and anharmonic oscillations.
  VOL. III. *Quantum Mechanics: Non-Relativistic Theory.* This includes discussion of the Periodic System, diatomic and multi-atomic molecules, the ideas of symmetry and the problem of the motion of an electron in a periodic field.
  VOL. V. *Statistical Physics.* This covers both classical and quantum statistics and includes specific heats of solids, thermal expansion, phonons, the quantum theory of liquids, superfluidity, the electron spectrum of solids, superconductivity, negative temperatures, equilibrium diagrams, the properties of solids at very high temperatures and pressures, isotopic mixtures, crystal symmetry, phase transitions, surface energies of crystals and crystal nucleation.
  VOL. VII. *Theory of Elasticity.* This includes crystal symmetry, elasticity and elastic waves in crystals as well as short wavelength effects, anharmonic oscillations and thermal conduction in crystals.
PEIERLS, R. *Quantum Theory of Solids.* Oxford Univ. Press 1955.
SEITZ, F. *The Modern Theory of Solids.* McGraw-Hill, New York and London 1940.

Chapter 1

SCHIFF, L. I. *Quantum Mechanics.* McGraw-Hill 1955.
SILLITTO, R. M. *Non-Relativistic Quantum Mechanics.* Edinburgh Univ. Press 1960.

Chapter 2

BORN, M. *Atomic Physics.* Blackie 1957.

Chapters 3 and 4

PAULING, L. *The Nature of the Chemical Bond and the Structure of Molecules and Crystals.* 3rd ed. Oxford Univ. Press 1963

Chapter 5

KITTEL, C. *Solid State Physics.* 2nd ed. Wiley, New York; Chapman and Hall, London 1956.

SPACE GROUPS

KOSTER, G. F. *Solid State Physics.* Vol. **5.** 1957.

SYMMETRY

NYE, J. F. *The Physical Properties of Crystals: their representation by tensors and matrices.* Oxford Univ. Press 1960.

SMITH, C. S. *Solid State Physics.* Vol. **6.** 1958.

Chapter 6

CRYSTAL CHEMISTRY

EVANS, R. C. *An Introduction to Crystal Chemistry.* 2nd ed. Cambridge Univ. Press 1964.

METALS AT HIGH TEMPERATURES AND PRESSURES

BUNDY, F. P. and H. M. STRANG. *Solid State Physics.* Vol. **13,** 1962.

SOLID SOLUTIONS

FRIEDEL, J. and A. GUINIER. *Metallic Solid Solutions.* Benjamin, New York 1963.

STRUCTURE OF LIQUID METALS

FRENKEL, J. *Kinetic Theory of Liquids.* Oxford Univ. Press 1946; Dover, New York 1960.

FROST, B. R. T. *Progress in Metal Physics.* Vol. **5,** 1954.

Chapter 7

COHESION

MOTT, N. F. *Rep. Progr. Physics.* Vol. **25,** 1962.

WIGNER, E. P. and F. SEITZ. *Solid State Physics.* Vol. **1,** 1955.

Chapter 8

WILSON, A. H. *The Theory of Metals.* Cambridge Univ. Press 1953.

FERMI SURFACE

PIPPARD, A. B. *Rep. Progr. Physics.* Vol. **23,** 1960.

ZIMAN, J. M. *Electrons in Metals: a Short Guide to the Fermi Surface.* Taylor and Francis, London 1963.

POSITRON ANNIHILATION

WALLACE, P. R. *Solid State Physics.* Vol. **10,** 1960.

ULTRASONIC ATTENUATION

MASON, W. P. *Physical Acoustics and the Properties of Solids.* Van Nostrand, New York 1958.

SEMICONDUCTORS

SMITH, R. A. *Semiconductors.* Cambridge Univ. Press 1959.

ELECTRON COMPOUNDS

HUME-ROTHERY, W. *Atomic Theory for Students of Metallurgy.* Inst. Metals London 1960.

Chapter 9

WILSON, A. H. *The Theory of Metals.* Cambridge Univ. Press 1953.

DIELECTRICS AND DIPOLE SUMS

HIPPEL, A. R. v. (Ed.) Dielectric Materials and Applications. Wiley, New York 1954.

ELECTRETS

GUTMAN, F. *Rev. Mod. Phys.* Vol. **20**, p. 457, 1948.

FERROELECTRICS AND ANTIFERROELECTRICS

DEVONSHIRE, A. F. *Rep. Progr. Physics.* Vol. **27**, 1964.
KANZIG, W. *Solid State Physics.* Vol. **4**, 1957.
MCGAW, H. D. *Ferroelectricity in Crystals.* Methven, London 1957.

ELECTRICAL CONDUCTIVITY

Encyclopaedia of Physics. Vol. **19**, Springer, Berlin 1956.
contains:
  GERRITSEN, A. N. *Metallic Conductivity*; (experimental part).
  JONES, H. *Theory of electrical and thermal conductivity in metals.*
  SLATER, J. C. *The Electronic Structure of Solids.*

SUPERCONDUCTIVITY

SCHAFROTH, M. R. *Solid State Physics.* Vol. **10**, 1960.
SCHOENBERG, D. *Superconductivity.* Cambridge Univ. Press 1952.

Chapter 10

MAGNETISM

BECKER, J. J. *Recent Developments in Magnetic Metals and Alloys. Metallurgical Reviews.* Vol. **7**, 1962.
BOZORTH, R. M. *Ferromagnetism.* Van Nostrand, New York 1951.

MAGNETIC DOMAINS

TEBBLE, R. S. and D. J. CRAIK. *Rep. Progr. Physics.* Vol. **24**, 1961.

FERRIMAGNETISM

WOLF, W. P. *Rep. Progr. Physics.* Vol. **24**, 1961.

NEUTRON DIFFRACTION STUDIES OF MAGNETIC STRUCTURE

BACON, G. E. *Neutron Diffraction.* Oxford Univ. Press 1962.

Chapter 11

PHASE CHANGES

TURNBULL, D. *Solid State Physics.* Vol. **3**, 1956.

Chapter 12

ORDER-DISORDER PHENOMENA IN METALS

GUTTMAN, L. *Solid State Physics.* Vol. **3**, 1956.

Chapter 13

GROWTH OF CRYSTALS FROM AQUEOUS SOLUTION

LAUDISE, R. A. and J. W. NIELSON. *Solid State Physics.* Vol. **12**, 1961.

GROWTH OF CRYSTALS BY PHASE CHANGE

TURNBULL, D. *Solid State Physics.* Vol. **3**, 1956.

Chapter 14

SPECIFIC HEATS

DE LAUNAY, J. *Solid State Physics.* Vol. **2**, 1955.
PARKINSON, D. H. *Rep. Progr. Physics.* Vol. **21**, 1958.

LATTICE VIBRATIONS

COCHRAN, W. *Rep. Progr. Physics.* Vol. **26**, 1963.
MITRA, S. S. *Solid State Physics.* Vol. **13**, 1962.

BeatriceGloria_personal library

THERMAL CONDUCTIVITY: (NON-METALS)

BERMAN, R. *Adv. in Physics.* Vol. **2**, p. 103, 1953.

THERMAL CONDUCTIVITY: (METALS)

KLEMENS, P. G. *Solid State Physics.* Vol. **7**, 1958.
MENDELSSOHN, K. and H. M. ROSENBERG. *Solid State Physics.* Vol. **12**, 1961.

Chapter 15

DIFFUSION

BARRER, R. M. *Diffusion in and through Solids.* Cambridge Univ. Press 1941.
JOST, W. *Diffusion in Solids, Liquids and Gases.* Academic Press, New York 1960.
LAZARUS, D. *Solid State Physics.* Vol. **10**, pp. 71-126, 1960.
LE CLAIRE, A. D. *Progress in Metal Physics* Vol. **1**, 1949; Vol. **4**, 1953.

POINT DEFECTS

Vacancies and other Point Defects in Metals and Alloys. Institute of Metals, London 1958.
LOMER, W. M. *Progress in Metal Physics.* Vol. **8**, 1959.

Chapter 16

COTTRELL, A. H. *The Mechanical Properties of Matter.* Wiley, New York 1964.
SCHMID, W. and G. BOAS. *Plasticity of Crystals.* F. A. Hughes, London 1950.

DISLOCATIONS

FRIEDEL, J. *Dislocations.* Pergamon Press 1964.
READ, W. T. *Dislocations in Crystals.* McGraw-Hill, New York 1953.

TWINNING

CLARK, R. and G. B. CRAIG. *Progress in Metal Physics.* Vol. **3**, 1952.

ELASTICITY

HUNTINGDON, H. B. *Solid State Physics.* Vol. **7**, 1958.

INDEX OF AUTHORS

Aggarwal, P., 168

Bardeen, J., 311
Barnett, S. J., 298, 325
Bassett, G. A., 458
Beevers, C. A., 295
Belov, N. V., 161, 194
Berman, R., 405
Bertaut, E. F., 179
Bethe, H., 353
Biermasz, T., 405
Blackman, M., 401
Bleik, W. E., 100
Bogolyubov, N. N., 311
Bohm, H. V., 309
Bohr, N., 64, 74, 88, 98
Bokii, G. B., 194
Boltzmann, L., 239, 335, 394
Born, M., 100, 206, 394, 397
Bouguer, P., 409
Bragg, W. L., 42, 194, 220, 223, 353
Bridgman, P. W., 449
Buckel, W., 168
Bykov, V. N., 184

Casimir, H. B. G., 411
Chatterjee, N. S., 168
Clausius, R., 249
Cooper, D., 311
Curie, P., 256ff, 322

Davidenkov, N. N., 452
Debye, P., 176, 252, 389
Dorfman, Ya. G., 327
Dulong, P. L., 228, 381

Eguchi, M., 235
Einstein, A., 325, 385, 387
Eremeev, N., 259
Evjen, H. M., 473
Ewald, P. P., 243

Faraday, M., 297, 318
Fedorov, E. S., 137, 152ff
Fermi, E., 64
Fersman, A. E., 207
Firsov, M. M., 441
Fock, V. A., 65
Foëx, G., 253
Frank, F. C., 374
Franz, R., 213, 300, 407
Frazer, B. C., 289, 295
Frenkel, Ya I., 425, 426, 454
Fröhlich, H., 311

Gadolin, A. V., 149
Gogoberidze, D. B., 441
Golder, G. A., 201, 202
Göler, F. von, 376, 483
Goldschmidt, V. M., 194, 201
Gorsky, V. S., 353
Goswani, A., 168
Grüneisen, E., 414
Gubar, C. V., 306

de Haas, W. J., 316, 325, 405
Haber, F., 206
Hall, E. H., 225
Hartree, D. R., 65
Hessel, J. F. C., 149
Hund, F., 66
Huntingdon, H. B., 128

Ingham, A. E., 200

Joel, N., 440
Jones, F. W., 364

Kammerlingh Onnes, H., 301
Kant, E., 76
Kapitza, P., 312
Kapustinsky, A. F., 207
Kármán, T. von, 394, 397
Keesom, W. H., 304
Khotkevich, V. I., 168
Kikoin, I. K., 306
Kitaigorodsky, A. I., 154
Kolmogorov, A. N., 378, 483
Konobeevsky, S. T., 428
Kopp, H., 381
Kornfeld, M. O., 434
Kurchatov, B. V., 259
Kurnakov, N., 350

Lambert, J. M., 409
Landau, L. D., 294, 304, 312, 316, 320
Langevin, P., 238, 322
Laue, M. von, 137
de Launay, B. N., 151
Laval, J., 439
Lemlein, G. G., 368
Lennard-Jones, J. E., 202
Lewis, G. N., 97
Le Corre, Y., 440
Livshits, I. M., 294, 345, 353
Lorentz, H. A., 213, 241, 250
Lorenz, L., 250, 407
Lukirsky, P. I., 369

McDonald, D. K. C., 406
McKeehan, L. W., 243

Madelung, E., 204, 208, 210
Mandelshtam, L. I., 298
Maxwell, J. C., 298
Mayer, J. E., 100
Meissner, W., 303, 310
Mendeleev, D. I., 68-75
Menter, J. W., 458
Meshkovsky, A. G., 304
Morse, R. W., 309
Moseley, H. G., 69
Mosotti, F. O., 249
Müller, E., 137

Néel, L., 332
Neumann, F. E., 381

Pashley, D. W., 458
Pauli, W., 97, 217
Pauling, L., 124, 194
Peierls, R., 405, 410
Pepinsky, R., 289
Petit, A. T., 228, 381

Quarrel, A., 168

Riecke, E., 298
Rivkin, M. M., 434

Sachs, G., 376, 483
Shal'nikov, A. I., 304
Shubnikov, A. V., 154
Simon, F. E., 405
Slater, J. C., 100
Sykes, C., 364

Takagi, Y., 282
Tarasov, V. V., 384
Taylor, G. I., 455
Thomas, L. H., 64

Ubbelohde, A. R., 129
Unsöld, A., 100

van Alphen, P. M., 316
van der Waals, J. D., 101, 106, 194
van Laer, P. H., 304
Vegard, L., 175
Venevtsev, Yu. N., 288
Volmer, M., 369
Vul, B. M., 261

Wiedemann, G., 213, 300, 407
Wigner, E. P., 128
Williams, E. J., 353
Wooster, W. A., 440
Wulf, Yu. V., 42, 138, 220
Wyckoff, R. W. G., 137, 368

Zagrubsky, A., 423, 431
Zhdanov, G. S., 288
Ziman, J. M., 405
Zubov, V. G., 441
Zwikker, C., 440

INDEX OF SUBJECTS

Amorphous state, 367
Anharmonicity:
  of thermal vibrations, 408, 411
  of zero-point oscillations, 409
Anisotropy, 136, 141
  of the dielectric constant, 247
  of elasticity, 435ff
  of electrical resistivity, 299
  magnetic, 328
  of thermal expansion, 412ff
Antiferroelectrics, 235, 258ff, 273ff
Antiferromagnetics, 315, 329
Atom of hydrogen: *see also* hydrogen atom
  energy levels, 58
  ionization potential, 58
  quantum number, 58
  Schrödinger equation, 55ff
  wave functions, 61
Atom, multi-electron:
  atomic nucleus-structure, 48
  electron binding energy, 83
  energy level scheme, 65
  order of filling of electron subgroups, 66
Atomic radii, 100, 168, 191ff, 209

Band theory of crystals, 223ff
Barkhausen effect, 326
Born-Haber cycle, 206
Born-Kármán theory of crystal vibrations, 394ff
Bohr atom, 64
Bohr molecule, 135, 470
Bond:
  chemical, 85, 108
  covalent, 108
  directed, 121
  donor-acceptor, 134
  heteropolar, 86, 129
  homopolar, 86, 108, 125
  hybrid, 121
  hydrogen, 105
  ionic, 86, 129ff
  metallic, 125
  multiple, 111
  quantum theory of, 89ff
  σ and π, 111
Bravais lattice, 148, 152
Brittleness:
  of crystals, 433, 441
  of liquids, 434

Characteristic temperature, 383, 387, 392
Chemical bond: *see* bond
Cleavage, 441ff
Close packing, 156ff
  numerical symbol of, 158ff
Cohesion, 197ff
  ionic crystals, 203ff, 210
  molecules, 198
  molecular crystals, 202
Compressibility:
  of ionic crystals, 205
  of metals, 210
Conductivity:
  electrolytic, 296ff
  electron, 233, 299ff
  hole, 233
  super-, 300ff
  thermal, 307, 402ff
Constant:
  Madelung, 203ff
  dielectric, 234, 247ff
Covalent bond, 88, 108
Crystallisation kinetics, 365ff
Crystal chemistry, 164ff
  diamond-like compounds, 188
  electron compounds, 185
  elements, 164
  ionic compounds, 192
  solid solutions, 173
Crystal symmetry classes, 141
Curie's law, 322
Curie temperature, 258, 322
Curie-Weiss law, 265, 267, 315, 322

Davidenkov's classification of internal stresses, 452ff
Debye theory of specific heat, 389ff
Defects, crystalline, 424ff
  dislocations, 368, 455ff
  interstitial atoms, 423ff
  vacancies, 423, 412
Deformation:
  elastic, 432, 434ff
  hot and cold, 452
  plastic, 433, 445ff
Diamagnetism, 313ff
  of atoms, 319
  of electron gas, 320
  of metals, 321
  universality of, 318
Dipole moment:
  of heteropolar molecules, 129
  of the water molecule, 133
Dielectric constant, 234, 247ff
Diffusion, 421ff
  atomic mechanism, 424ff
  activation energy for, 424ff
  in electrolytes, 429
  measurement of, 423
  surface, 369

Diffusion (*cont.*)
  uphill, 427
  volume, 421
Dislocations:
  edge (linear), 455
  formation of, 458
  screw, 374
Dispersion forces, 101
Dispersion of elastic waves in crystals, 394ff
Distribution of elements, 50ff
Domains:
  ferroelectric, 294
  ferromagnetic, 325
Donor-acceptor bond, 134
Dulong and Petit's law, 214, 381

Effect:
  Barkhausen, 326
  de Haas-van Alphen, 316
  Hall, 45, 225
Effective mass, 44
Einstein-de Haas experiment, 325
Einstein's theory of specific heat, 384
Elasticity of crystals, 434ff
  Cauchy's relations, 438
  generalised theory, 440
  strain tensor, 437
  stress tensor, 435
Electrets, 235, 253
Electron compounds:
  in crystal chemistry, 185
  theory, 229ff
Electron diffraction, 168
Electron micrographs of molecular structures, 137, 368, 458
Electron theory of metals:
  classical, 212
  quantum, 217ff
Electrical resistivity:
  anisotropy of, 299
  dependence on deformation concentration and temperature, 300
  of metals, 212, 298
  of solid electrolytes, 296
Elementary particles, 47
Elements, chemical, periodicity of properties, 69ff
  electron structure, 76
Energy:
  activation, 423
  binding energy of electron, 83
  cohesive, 197ff
  free, 334
  internal, 333
  of internal stresses, 452
  of a quantum oscillator, 384
Energy bands, 223ff
Entropy, 334
Equilibrium diagrams, 342ff
Expansion: *see* thermal expansion
Extrinsic conductivity, 233

Fedorov groups, 152
Ferrites, 330
Ferroelectrics, 50, 235
Ferromagnetism, 325ff
Fluctuations, 372
Formation of crystals, 365ff
Free energy, 334

Glassy state, 367
Goldschmid table of ionic radii, 192
Groups:
  Fedorov, 152
  point, 144
  Shubnikov, 154
  space, 144
Growth of crystals, 365ff
Gyromagnetic anomaly, 325
Gyromagnetic ratio:
  orbital, 60, 318
  spin, 47

Hall constant, 45, 225
Hamiltonian operator, 1ff
Hardening:
  alloy, 451
  thermal, 450
  orientation, 449
  pressure, 450
  work, 451
Heat of sublimation:
  of molecular crystals, 203
  of the elements, 464
Heteropolar bond, 86, 129ff
Heusler alloys, 177, 317, 318
Hill's equation, 42
Holes, 225, 233
Homopolar bond, 86, 108, 125
Hooke's law, 408, 434ff
Hund's rules, 66, 327
Hybridisation of electronic states, 6, 121
Hydrogen: *see also* atom of hydrogen, 55ff
  bond, 105
  molecule, 111
  molecular ion, 108
  wave functions, 61

Insulators, 224, 234
Interstitials, 423ff
Internal energy of a system, 333
Internal potential of a crystal, 214
Internal stresses:
  classification, 452ff
  energy, 452
Ionic binding, 86, 129, 203ff
Ionic model of a metal, 208
Ion microscope, resolution of atomic structure, 137
Ionic radii—table, 192
Ionisation potential of atoms, 82
Isotopes, 49
  influence on the physical properties of crystals, 49
Isotopic effects in crystals, 50, 101

Kolmogorov's theory of the kinetics of crystallisation, 378
Kurnakov temperature, 350; *see also* temperature, ordering

Latent heat of crystallisation, 371
Landau's theory of diamagnetism of free electrons, 320
Laplacian operator, 3
Lattice, Bravais, 148, 152
Laws:
  Bouguer, 409
  Curie-Weiss (dielectric), 267, (magnetic), 322
  Dulong and Petit, 214, 381
  Fick, 422
  Goldschmid, 192
  Gruneisen, 414
  Hooke, 408, 417, 434ff
  Lambert, 409
  Lorenz-Lorentz, 250
  Mendeleev, 68
  Neumann-Kopp, 261, 381
  Ohm, 213, 298
  Planck, 29
  Stefan-Boltzmann, 394
  Vegard, 175
  Weidemann-Franz, 213, 407, 411
Legendre's equation, 33
Long range order, 352, 353ff

Madelung constant, 204
Magnetic moment:
  of atoms of transition metals, 328
  of the electron, neutron and proton, 47
  orbital, 59
Magnetic structure, 329
Magnetic susceptibility, 313
  of the elements, 467
Magnetisation, 313
Matthieu's equation, 40
Mendeleev's law, 68ff; *see* periodic system
Metals, 74, 233
  alkali and alkaline earth, 76
  rare earth, 79
  transition, 77
Metalloids, 85
  position in the periodic table, 74
Metastable phases, 349
Mobility of ions in solid electrolytes, 296ff, 429ff
Molecule:
  diatomic, 111
  hydrogen, 111
  lithium, 111
  polyatomic, 119
  oxygen, 116
  water, 120
Molecular ion of hydrogen, 108
Morphotropy, 189

Neutron spectroscopy, 284, 329
Nucleus:
  atomic, 48
  of crystallisation, 365ff
  two dimensional, 373

Octet rule, 170
Order, 350ff
  long range, 352, 353ff
  of a solid solution, 177
  short range, 353, 361ff

Packing fraction, 161
Paramagnetism, 313, 321ff
  at low temperatures, 324
  Langevin theory, 321
  of the electron gas, 322
  of the transition metals, 323
  Weiss theory, 322
Partition function, 385
Pauling's table of ionic radii, 192
Periodic system, 68ff, 74ff
  properties of elements, 461ff
Permeability, 314
Phase equilibrium, 333ff
Phase transitions, 347ff
  eutectic, 348
  first and second order, 351
Phonons, 307, 394
Piezoelectrics, 235, 253ff
Planck's law, 29
Plane:
  glide, 140
  slip, 446
  twinning, 459
Plasticity of crystals, 445ff
  critical stress for, 448
  slip, 445
  theory, 455
Polarisability, 235ff
Polymorphism, 164ff
Polytypism, 161, 189, 368
Positron annihilation, 216

Quantum numbers, 58

Radii:
  atomic, 168ff, 192
  correction for coordination, 192, 201
  correction for multiplicity of bond, 192
  ionic, 192, 195
  molecular, 170, 192
Rose, fracture, 441

Scattering of electron waves in a crystal, 220
Schrödinger equation:
  electron in a periodic field, 39
  free particle, 10ff
  for central forces, 30
  harmonic oscillator, 27ff
  hydrogen atom, 55, 108
  particles in a potential well, 20ff
  rigid rotator, 46
  time dependent, 2, 7
  time independent, 4
  variational method for solution, 92
Self-diffusion: *see* diffusion
Semiconductors:
  energy schemes, 225
  n- and p-type, 233
Slip, 445ff
  systems, 446
Solid solutions:
  distorted structures, 175
  equilibrium diagrams, 342ff
  interstitial, 184
  ordered, 177, 342ff
  substitutional, 174

Solubility, 340ff
Space groups, 144, 152
  distribution of crystals among, 153
Space lattice, 140
Specific heat, 380ff
  and structure, 383
  classical, 381
  molecular crystals, 384
  of the electron gas, 219
  quantum theory of, 384ff
  temperature dependence, 382
Spiral growth of crystals, 361, 368
Spontaneous:
  magnetisation, 325
  polarisation, 294
Stoichiometry, 131
Strength:
  theoretical, 208, 442
  practical, 208
Structural coefficients, 243
Structural sums, 200
Structures:
  binary compounds, 185
  chain type, 156, 170
  close-packed spherical, 156
  coordination, 156, 164
  diamond-like compounds, 188
  electron compounds, 185
  elements, 164
  ionic compounds, 191
  layer type, 156, 170
  molecular, 155
  solid solutions, 173ff
  with directed bonds, 155
Sum of states, 385; *see also* partition function
Superconductivity, 168, 300ff
  theory, 311
Superperiodicity, 161, 189
Symmetry, 136ff
  of atoms and molecules, 145
  of crystals, 144
  of elements, 138
  operations, 136
Symmetry axes, 138
Symmetry groups:
  point, 144ff
  space, 144, 152ff

Temperature:
  Curie (dielectric), 258
  Curie (magnetic), 322
  Debye (characteristic), 383
  Kurnakov, 350ff
  ordering, 350ff
Tensor:
  deformation, 437
  elasticity, 437
  stress, 436
  thermal expansion, 412
Tensile strength, 441
Thermal conductivity, 402ff
  influence of neutron irradiation, 404
  of metals, 405
  of non metals, 402
  theory of, 407
  and Umklapp processes, 410
Thermal expansion, 411ff
  and crystal structure, 414
  temperature dependence, 414
  tensor, 412
  theory, 417ff
Translation operators, 136ff
Tunnel effect, 19
Twinning, 458ff

Umklapp processes, 410
Unit cell, 149

Vacancies, 421, 425
Van-der-Waal's forces, 101
Virial theorem, 472

Wiedemann-Franz law, 213, 407, 411
Work function, 215
Work hardening, 451ff

X-ray spectrum, 126

Zero point energy of an oscillator, 29
  of isotopic crystals, 50, 101